

82178usprov

Expression vector, methods for the production of heterologous gene products and for the selection of recombinant cells producing high levels of such products

Scope of the Invention

The invention relates to a method for selecting highly productive recombinant cells, a method for preparing heterologous gene products and expression vectors and host cells transfected therewith which may be used in these processes.

Background to the Invention

Mammalian cells are the preferred host cells for the production of complex biopharmaceutical proteins as the modifications carried out post-translationally are compatible with humans both functionally and pharmacokinetically. Commercially relevant cell types are hybridoma, myeloma CHO (Chinese Hamster Ovary) cells and BHK (Baby Hamster Kidney) cells. The cultivation of the host cells is increasingly carried out under serum- and protein-free production conditions. The reasons for these are the concomitant cost reduction, the reduced interference in the purification of the recombinant protein and the reduction in the potential for the introduction of pathogens (e.g. prions and viruses). The use of CHO cells as host cells is becoming more widespread as these cells adapt to suspension growth in serum- and protein-free medium and are also regarded and accepted as safe production cells by the regulatory authorities.

In order to produce a stable mammalian cell line which expresses a heterologous gene of interest, the heterologous gene is generally inserted in the desired cell line together with a selectable marker gene such as e.g. neomycin phosphotransferase by transfection. The heterologous gene and the selectable marker gene can be expressed either together by a single

vector or by separate vectors which are cotransfected. Two to three days after transfection the cells are transferred into medium containing a selective agent, e.g. G418 when using neomycin phosphotransferase-gene, and cultivated for some weeks under these selective conditions. The emergent resistance cells can then be isolated and investigated for expression of the desired gene product. As a result of the random and undirected integration into the host cell genome a population of cells is obtained which have completely different rates of expression of the heterologous gene. These may also include non-expressing cells in which the selectable marker is expressed but not the gene of interest. In order to identify cell clones which have a very high expression of the heterologous gene of interest, it is therefore necessary to examine and test a large number of clones, which is time consuming, labour intensive and expensive.

Gene amplification is a widespread phenomenon in animal cell cultures, which is used for the production of recombinant biopharmaceutical proteins. The gene amplification drastically improves the originally relatively low productivity of numerous mammalian cell lines. One amplification technique which is widely used is dihydrofolate reductase (DHFR)-based gene amplification system which is very often used in DHFR-deficient Chinese Hamster Ovary (CHO) cells. The DHFR-deficient CHO cells, e.g. CHO-DUKX (ATCC CRL-9096) or CHO-DG44 (Urlaub et al., 1983), are transfected with a suitable vector system which codes for DHFR and the protein of interest. Then the transfectants are selected in a medium without glycine, hypoxanthine and thymidine. The amplification and hence the establishment of highly productive cell lines is achieved by the increasing addition of methotrexate (MTX), an inhibitor of dihydrofolate reductase (Kaufman et al., 1982; US 4,656,134). The subsequent selection of the highly productive cells obtained is also subject to the principle of chance and is based on probabilities, as a result of which this selection step is highly labour-intensive and time-consuming.

All kinds of methods have been developed for monitoring gene transformation and expression better and more rapidly. These include, first of all, the use of

reporter molecules such as chloramphenicol-acetyltransferase, luciferase, β -galactosidase or fusion proteins which contain the coding regions of β -galactosidase or luciferase. The disadvantage of these reporter gene assays is that the cells have to be fixed or lysed and have to be incubated with exogenously added substrates and co-factors. Thus, further cultivation of the analysed cells is out of the question. A more recent method based on the co-expression of the *E.coli* enzyme β -galactosidase does indeed allow lysed cells to be sorted using a FACS apparatus (Nolan et al., 1988), but hypotonic pretreatment is required in order to charge the cells with the fluorogenic substrate. This activity also has to be inhibited before the FACS-based sorting.

With the introduction of green fluorescent protein (GFP) from *Aequorea victoria* and the GFP mutants developed therefrom as reporter molecule it became much easier to identify cells which express a heterologous gene. Co-expression of GFP allowed real-time analysis in vivo and sorting of transfecteds on the basis of their fluorescence without the need for additional substrates or co-factors. The use of GFP as a reporter molecule for monitoring gene transfer has been described in various publications. In US patents 5,491,084 and 6,146,826, Chalfie et al. described a method of selecting cells which express a protein of interest. This method comprises co-transfection of cells by a DNA molecule which contains the coding sequence for the protein of interest, and a second DNA-molecule which codes the GFP-gene. Then the GFP-expressing cells are selected. Gubin et al. (1997) investigated the stability of GFP expression in CHO cells in the absence of selective growth conditions. The cells were transfected with a plasmid which contained both GFP and neomycin phosphotransferase. Mosser et al. (1997) used a plasmid which contained a bicistronic expression cassette coding for a GFP and a target gene (also known as the gene of interest) to identify and select cells which expressed inducible product. The target gene was under the control of a regulatable promoter. The coupling of the GFP and target gene expression was achieved using a viral IRES (Internal Ribosome Entry Site) element, as a result of which a bicistronic

mRNA which coded for GFP and the protein of interest was expressed. The plasmid used did not itself contain any selectable marker gene. This was therefore introduced by a second plasmid in a co-transfection or in a subsequent transfection. By contrast, Levenson et al. (1998) used retroviral vectors with a bicistronic expression cassette in which the gene of interest can be cloned in front of the IRES sequence. The sequence following the IRES sequence, on the other hand, coded for a selectable marker gene, this being a marker which conferred resistance to G418, puromycin, hygromycin B, histidinol D or phleomycin, or it was GFP.

Vectors have already also been described which contain an IRES element from the family of the picorna viruses, the IRES element being positioned between the product gene and a selectable marker gene (Pelletier et al., 1988; Jang et al., 1989; Davies et al., 1992).

GFP has also been successfully fused with resistance marker genes. For example, Bennett et al. (1998) describe a GFP/zeomycin fusion protein. This bifunctional selectable marker was successfully used to identify and select transfected mammalian cells. Primig et al. (1998) on the other hand used a fusion protein of GFP and neomycin phosphotransferase for their enhancer studies.

In the publication by Meng et al. (2000) and in International Patent application WO 01/04306, an expression system in which the gene of interest was expressed together with the amplifiable selectable marker gene DHFR and a GFP gene from a single vector was used to select and identify cells with a high expression of a recombinant protein. The three genes were either combined in one transcription unit or divided between two units. This spatial and transcriptional linking of all three genes in a single expression vector was intended to increase their probability of co-amplification under selection pressure and thus identify and select high producing clones. The best clones which were isolated by using the combined selection by means of amplifiable DHFR selection markers and GFP-based FACS sorting expressed the protein of interest in an order of magnitude of not more than 3 to 4.5 pg per cell per

day. The experiments were carried out with adherent cells and in serum-containing medium, i.e. with cells and under conditions which are known to be substantially more robust and are characterised by higher basic productivities.

Summary of the Invention

The aim of the present invention was therefore to develop a selection system for recombinant cells with increased productivity, which meets the following requirements:

- (1) A reduction in the time taken to develop high producing cells to produce biopharmaceutical proteins while at the same time lowering the development costs;
- (2) A high throughput in the selection of high producing cells with low expenditure on capacity;
- (3) The use of "fermentation-robust" high-producing cells which exhibit, for example, lower impairment of growth at increased methotrexate concentrations;
- (4) The transfection, selection and cultivation of the suspension-adapted cells, preferably in serum-free medium;
- (5) A reduction in the gene amplification steps required.

A further aim of the invention was to provide expression vectors and host cells transfected therewith which can be used in this clone selection system, as well as a process for preparing heterologous gene products using these host cells.

These objectives are achieved according to one aspect of the present invention by means of an expression vector which comprises a gene coding for a protein of interest (hereinafter also referred to as the "gene of interest") functionally linked to a hamster ubiquitin/S27a promoter and a gene which codes for a fluorescent protein.

The expression vector preferably also contains an amplifiable selection marker gene, e.g. the gene for dihydrofolate reductase (DHFR). A preferred expression vector also contains other regulatory elements, e.g. an enhancer functionally linked to the promoter. In addition, the expression vector preferably also contains an internal ribosomal entry site (IRES) which allows bicistronic expression of the gene which codes for a fluorescent protein and of the gene of interest.

The invention also relates to basic vectors which instead of the gene of interest have a multiple cloning site for the incorporation of such a gene, i.e. a sequence area with multiple cutting sites for restriction endonucleases.

In another aspect the present invention relates to host cells which have been transfected with one of the expression vectors mentioned. These are eukaryotic host cells, preferably mammalian cells, while rodent cells such as hamster cells and especially CHO cells or BHK cells are particularly preferred.

In another aspect the present invention relates to a process for preparing a heterologous gene product in which a host cell transfected with the expression vector according to the invention is cultivated under conditions which allow expression of the gene product and the gene product is isolated from the culture or the culture medium.

In one particular embodiment of the invention the host cell is transfected, preferably co-transfected, with the expression vector according to the invention and additionally with one or more vectors with genes which code for one or more other proteins of interest.

In this connection the present invention provides a process for preparing a heterodimeric protein in which a host cell of this kind which has been co-transfected with expression vectors which code for different sub-units of the heterodimeric protein is cultivated under conditions which allow expression of the heterodimeric protein, and the heterodimeric protein is isolated from the

culture or culture medium. One particular application for such a process is the production of antibodies and their sub-units.

In another aspect the present invention relates to a process for selecting a host cell which expresses a protein of interest in which a population of host cells which have been transfected with an expression vector according to the invention is cultivated under conditions which allow expression of the protein of interest and of the fluorescent protein, and the cell or cells which exhibit the highest expression rates of fluorescent protein are identified and/or selected. The selection is preferably made using a fluorescence-activated cell sorter (FACS).

Surprisingly, it has been found that using the system provided according to the invention it is possible in a very short time to isolate cell pools which express average specific productivities of more than 15 pg (single-chained protein) or 10 pg (humanised antibody) of recombinant protein per cell and per day, without a gene amplification step. The specific productivities could be increased to more than 30 pg per cell and per day by a single DHFR-based gene amplification step. The productivities achieved in the cell pools are hence higher than the maximum productivities of the best cell clones published hitherto by a factor of 8 to 10.

Astonishingly, there is also a very good correlation between the expression of the protein of interest and the fluorescent protein. This is even true in the case of co-transfection if - as with an expressed antibody - the two immunoglobulin chains are each expressed by their own vector and in the FACS sorting a selection can only be made for the expression of the one chain, on account of its transcriptional coupling to the fluorescent protein. The high expression rates of the fluorescent protein have no negative effect whatsoever on the growth and vitality of the cells. In addition, the development time for selecting high producing cells can be reduced by at least half compared with a conventional stepwise gene amplification strategy, resulting in a significant reduction in the development capacity and costs.

Description of the Figures

Figure 1 shows a comparison of the expression levels obtained with recombinant cell clones in which the heterologous gene product is expressed either under the control of the CMV promoter or under the control of a hamster-ubiquitin/S27a promoter. The two promoters are functionally linked to the CMV enhancer and the termination sequence, BGH poly A, is identical in every case. In the case of CMV¹ the expression vector is pcDNA3-based (Invitrogen, Kalsruhe, DE), in CMV² it is a pBluescript-based expression vector (Stratagene, La Jolla, CA, US) and in the case of CHO it is a pAD-CMV-based expression vector (Werner et al., 1998). In the expression of the lysosomal enzyme, all the expression vectors contain the amplifiable selectable marker dihydrofolate reductase (DHFR) and the expression of the heterologous gene has been increased by subsequent amplification steps with methotrexate (MTX) . In order to express the two chains of the antibody (Ab) co-transfection has been carried out with a second vector which contains a neomycin-resistance gene as the selectable marker. The titres or specific productivities obtained are given in relation to the CMV promoter-based expression, which is set at 1 (CMV¹ for enzyme, CMV² for Ab).

Figure 2 shows a diagrammatic representation of the base vectors used to express the recombinant proteins in CHO-DG44 cells. "P/E" is a combination of CMV enhancer and hamster-ubiquitin/S27a promoter, "P" on its own indicates a promoter element and "T" is a termination signal for transcription, which is needed for the polyadenylation of the transcribed mRNA. The position and direction of transcription initiation within each transcription unit is indicated by an arrow. For cloning the heterologous genes a sequence region with multiple cutting sites for restriction endonucleases (multiple cloning sites - mcs) is inserted after the promoter element. The amplifiable selectable marker dihydrofolate reductase is abbreviated to "dhfr" and the selectable marker neomycin phosphotransferase is abbreviated to "neo". The "IRES" element coming from the encephalomyocarditic virus acts as an internal ribosomal entry site within the bicistronic transcription unit and enables translation of the following green fluorescent protein "GFP".

Figure 3 is a diagrammatic representation of the eukaryotic expression vectors each of which codes for a biopharmaceutical protein and has been used to transfect CHO-DG44 cells. "P/E" is a combination of CMV enhancer and hamster-ubiquitin/S27a promoter, "P" on its own indicates a promoter element and "T" is a termination signal for transcription, which is needed for the polyadenylation of the transcribed mRNA. The position and direction of transcription initiation within each transcription unit is indicated by an arrow. The amplifiable selectable marker dihydrofolate reductase is abbreviated to "dhfr" and the selectable marker neomycin phosphotransferase is abbreviated to "neo". The "IRES" element coming from the encephalomyocarditis virus acts as an internal ribosomal entry site within the bicistronic transcription unit and enables translation of the following green fluorescent protein "GFP". "sICAM" codes for the soluble intracellular adhesion molecule (US 5,412,216), whereas "F19HC" and "F19LC" code for the heavy and light chains, respectively of the humanised antibody F19 (EP 953 639).

Figure 4 shows the correlation between the sICAM productivity and the GFP fluorescence taking the cell pool ZB1 as an example. This cell pool was obtained from the transfection with the vector pBIDG-sICAM, in which the therapeutic protein sICAM and GFP are jointly expressed by a bicistronic transcription unit. The pool was subjected to a sequential GFP-based FACS sorting. After each sorting step (sort) the concentration of the sICAMs in the cell culture supernatant of the pool was determined by ELISA and the specific productivity per cell and per day was calculated (pg/c*d). Each data point is the average of at least three cultivation runs. A total of six sorts were carried out.

Figure 5 shows the isolation of high-expressing sICAM cells by GFP-based FACS sorting taking the cell pool ZB1 as an example. This cell pool was obtained from the transfection with the vector pBIDG-sICAM in which the therapeutic protein sICAM and GFP are together expressed by a bicistronic transcription unit. The pool was subjected to sequential GFP-based FACS sorting. After each sort the concentration of the sICAM in the cell culture

supernatant of the pool was determined by ELISA and the specific productivity per cell and per day (pg/c*d) was calculated. Each datapoint represents the average of at least three cultivation runs. In all, six sorts were carried out.

Figure 6 shows the increase in sICAM productivity achieved by combining GFP-based selection with an MTX amplification step, taking the cell pool ZB1 as an example. This cell pool, which was obtained from the transfection with the vector pBIDG-sICAM, was subjected to sequential GFP-based FACS sorting. After the fourth sort and sixth sort a DHFR-mediated gene amplification was carried out by adding methotrexate (MTX) to the cultivation medium (5 nM, 50 nM, 500 nM or 2 μ M MTX). The concentration of the sICAMs in the cell culture supernatant of the pool was determined by ELISA and the specific productivity per cell and per day (pg/c*d) was calculated. Each datapoint represents the average of at least three cultivation runs. Figure 7 shows the viability pattern of cell pools after the addition of different doses of methotrexate to the cultivation medium. The cell pool ZB1 which was obtained by transfection with the vector pBIDG-sICAM (Fig.3) was subjected to sequential GFP-based FACS sorting. After the fourth sort and sixth sort a DHFR-mediated gene amplification was carried out by adding methotrexate (MTX) to the cultivation medium. The cell numbers and viability were determined during the selection phase by staining with tryptan blue and monitored over a number of days in cultivation (dic).

Figure 8 shows the correlation between the antibody productivity (mAb F19) and the GFP fluorescence taking the cell pool ZB1 as an example. This cell pool was obtained from the transfection with the vector combination pBIDG-F19HC and pBIN-F19LC (Fig.3). The pool was subjected to sequential GFP-based FACS sorting. After each sort the concentration of the antibody F19 in the cell culture supernatant of the pool was determined by ELISA and the specific productivity per cell and per day (pg/c*d) was calculated. Each datapoint represents the average of at least three cultivation runs. In all, six sorts were carried out.

Figure 9 shows the isolation of high expressing mAbF19 cellpools by a GFP-based selection using FACS taking the cell pool ZB1 as an example. This cell pool, which was obtained by co-transfection with the vectors pBIDG-F19HC and pBIN-F19LC (Fig.3), was subjected to sequential GFP-based FACS sorting. The concentration of the antibody F19 in the cell culture supernatant of the pool was determined by ELISA after each sort and the specific productivity per cell and per day (pg/c*d) was calculated. Each datapoint represents the average of at least three cultivation runs.

Detailed description of the invention and preferred embodiments

The expression vector according to the invention contains a gene which codes for a protein of interest ("gene of interest "), functionally linked to a hamster ubiquitin/S27a promoter and a gene which codes for a fluorescent protein. Preferably, the expression vector also contains an amplifiable selectable marker gene.

Hamster-Ubiquitin/S27a Promoter

The ubiquitin/S27a promoter of the hamster is a powerful homologous promoter which is described in WO 97/15664. Such a promoter preferably has at least one of the following features: GC-rich sequence area, Sp1 binding site, polypyrimidine element, absence of a TATA box. Particularly preferred is a promoter which has an Sp1 binding site but no TATA box. Also preferred is a promoter which is constitutively activated and in particular is equally active under serum-containing, low-serum and serum-free cell culture conditions. In another embodiment it is an inducible promoter, particularly a promoter which is activated by the removal of serum.

A particularly advantageous embodiment is a promoter with a nucleotide sequence as contained in Fig. 5 of WO 97/15664. Particularly preferred are promoter sequences which contain the sequence from position -161 to -45 of Fig. 5.

The promoters used in the examples of the present patent specification each contain a DNA molecule with the sequence from position 1923 to 2406 of SEQ ID NO:1 of the attached sequence listing. This sequence corresponds to the fragment -372 to +111 from Fig. 5 of WO 97/15664 and represents the preferred promoter, i.e a preferred promoter should incorporate this sequence region. Another suitable promoter fragment contains the sequence from position 2134 to 2406 (corresponding to -161 to +111 in Fig. 5 of WO 97/15664). A promoter which contains only the sequence from position 2251 to 2406 is no longer functional (corresponds to position -45 to +111 in Fig. 5 of WO 97/15664). It is possible to extend the promoter sequence in the 5' direction starting from position 2134.

It is also possible to use functional subfragments of the complete hamster ubiquitin/S27a promoter sequence as well as functional mutants/variants of the complete sequence of subfragments thereof which have been modified, for example, by substitution, insertion or deletion. Corresponding subfragments, mutants or variants are hereinafter also referred to as "modified promoters".

A modified promoter, optionally combined with other regulatory elements, preferably has a transcription activity which corresponds to that of the promoter fragment from position 1923 to 2406 of the nucleotide sequence given in SEQ ID NO:1 (-372 to +111 from Fig. 5 of WO 97/15664). A modified promoter proves to be useful for the purposes of the invention if it has a transcription activity which has at least 50%, preferably at least 80%, more preferably at least 90% and most preferably at least 100% of the activity of the 1923 to 2406 fragment (-372 to +111 fragment) in a comparative reporter gene assay. Particularly preferred are modified promoters which have a minimum sequence homology to the wild-type sequence SEQ ID NO:1 of the hamster ubiquitin/ S27a promoter of at least 80%, preferably at least 85%, preferably at least 90%, more preferably at least 95% and most preferably at least 97% and have a corresponding promoter activity in a comparative reporter gene assay.

In a corresponding comparative reporter gene assay the promoter fragments to be tested including the reference sequence are cloned in front of a promoterless reporter gene which codes, for example for luciferase, secreted alkaline phosphatase or green fluorescent protein (GFP). These constructs (promoter sequence + reporter gene) are subsequently introduced into the test cells, e.g. CHO-DG44, by transfection and the induction of the reporter gene expression by the promoter fragment in question is determined by measuring the protein content of the reporter gene. A corresponding test is found for example in Ausubel et al., Current Protocols in Molecular Biology, 1994, updated.

The promoter sequence of the hamster ubiquitin/S27a promoter and the modified promoters, which may also include, for example, the 5' untranslated region or selected fragments thereof, and the coding region of the ubiquitin/S27a gene or selected fragments thereof, may be obtained by a skilled man with a knowledge of the sequence described in WO 97/15664 using various standard methods as described for example in Sambrook et al., 1989; Ausubel et al., 1994. Starting from the sequence described in WO 97/15664 a suitable fragment may be selected, for example, and an oligonucleotide probe containing the sequence of this fraction may be chemically synthesised. A probe of this kind may be used for example to clone the ubiquitin/S27a gene or the 5' untranslated region or other fragments thereof, for example by hybridisation from a library of the hamster genome. Using the reporter gene assay described above the skilled man is in a position to identify promoter-active fragments without any great effort and use them for the purposes of the present invention. The 5' untranslated region or special fragments thereof can easily be obtained by PCR amplification with corresponding primers from genomic DNA or a genomic library. Fragments of the 5' untranslated region may also be obtained by limited exonuclease III digestion from larger DNA fragments. Such DNA molecules may also be chemically synthesised or produced from chemically synthesised fragments by ligation.

Deletion, insertion and substitution mutants may be produced by "site-specific mutagenesis" and/or "PCR-based mutagenesis techniques". Corresponding methods are mentioned for example in Lottspeich and Zorbas 1998 Chapter 36.1 with further references.

By cross-hybridisation with probes from the 5' untranslated region of the hamster ubiquitin/S27a gene or from the S27a part of the hamster ubiquitin S27a gene it is also possible to identify and isolate suitable promoter sequences from corresponding homologous genes of other, preferably mammalian species. Suitable techniques are described by way of example in Lottspeich and Zorbas 1998 Chapter 23. Genes are "homologous" for the purposes of the invention if their nucleotide sequence exhibits at least 70%, preferably at least 80%, preferably at least 90%, more preferably at least 95% and most preferably at least 97% conformity to the nucleotide sequence of the gene with which it is homologous.

Gene of Interest

The gene of interest contained in the expression vector according to the invention comprises a nucleotide sequence of any length which codes for a product of interest. The gene product or "product of interest" is generally a protein, polypeptide, peptide or fragment or derivative thereof. However, it may also be RNA or antisense RNA. The gene of interest may be present in its full length, in shortened form, as a fusion gene or as a labelled gene. It may be genomic DNA or preferably cDNA or corresponding fragments or fusions. The gene of interest may be the native gene sequence, or it may be mutated or otherwise modified. Such modifications include codon optimisations for adapting to a particular host cell and humanisation. The gene of interest may, for example, code for a secreted, cytoplasmic, nuclear-located, membrane-bound or cell surface-bound polypeptide.

The term "nucleotide sequence" or "nucleic acid sequence" indicates an oligonucleotide, nucleotides, polynucleotides and fragments thereof as well as DNA or RNA of genomic or synthetic origin which occur as single or double

strands and can represent the coding or non-coding strand of a gene. Nucleic acid sequences may be modified using standard techniques such as site-specific mutagenesis or PCR-mediated mutagenesis (e.g. described in Sambrook et al., 1989 or Ausubel et al., 1994).

By "coding" is meant the property or capacity of a specific sequence of nucleotides in a nucleic acid, for example a gene in a chromosome or an mRNA, to act as a matrix for the synthesis of other polymers and macromolecules such as for example rRNA, tRNA, mRNA, other RNA molecules, cDNA or polypeptides in a biological process. Accordingly, a gene codes for a protein if the desired protein is produced in a cell or another biological system by transcription and subsequent translation of the mRNA. Both the coding strand whose nucleotide sequence is identical to the mRNA sequence and is normally also given in sequence databanks, e.g. EMBL or GenBank, and also the non-coding strand of a gene or cDNA which acts as the matrix for transcription may be referred to as coding for a product or protein. A nucleic acid which codes for a protein also includes nucleic acids which have a different order of nucleotide sequence on the basis of the degenerate genetic code but result in the same amino acid sequence of the protein. Nucleic acid sequences which code for proteins may also contain introns.

The term cDNA denotes deoxyribonucleic acids which are prepared by reverse transcription and synthesis of the second DNA strand from a mRNA or other RNA produced from a gene. If the cDNA is present as a double stranded DNA molecule it contains both a coding and a non-coding strand.

The term intron denotes non-coding nucleotide sequences of any length. They occur naturally in numerous eukaryotic genes and are eliminated from a previously transcribed mRNA precursor by a process known as splicing. This requires precise excision of the intron at the 5' and 3' ends and correct joining of the resulting mRNA ends so as to produce a mature processed mRNA with the correct reading frame for successful protein synthesis. Many of the splice donor and splice acceptor sites involved in this splicing process, i.e. the

sequences located directly at the exon-intron or intron-exon interfaces, have been characterised. For an overview see Ohshima et al., 1987.

Protein of Interest

Proteins/polypeptides with a biopharmaceutical significance include for example antibodies, enzymes, cytokines, lymphokines, adhesion molecules, receptors and the derivatives or fragments thereof, but are not restricted thereto. Generally, all polypeptides which act as agonists or antagonists and/or have therapeutic or diagnostic applications are of value.

The term "polypeptides" is used for amino acid sequences or proteins and refers to polymers of amino acids of any length. This term also includes proteins which have been modified post-translationally by reactions such as glycosylation, phosphorylation, acetylation or protein processing. The structure of the polypeptide may be modified, for example, by substitution, deletion or insertion of amino acids and fusion with other proteins while retaining its biological activity.

Examples of therapeutic proteins are insulin, insulin-like growth factor, human growth hormone (hGH) and other growth factors, tissue plasminogen activator (tPA), erythropoietin (EPO), cytokines, e.g. interleukines (IL) such as IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-11, IL-12, IL-13, IL-14, IL-15, IL-16, IL-17, IL-18 interferon (IFN)-alpha, -beta, -gamma, -omega or -tau, tumour necrosis factor (TNF) such as TNF-alpha, beta or gamma, TRAIL, G-CSF, GM-CSF, M-CSF, MCP-1 and VEGF. Other examples are monoclonal, polyclonal, multispecific and single chain antibodies and fragments thereof such as for example Fab, Fab', F(ab')₂, Fc and Fc' fragments, light (L) and heavy (H) immunoglobulin chains and the constant, variable or hypervariable regions thereof as well as Fv and Fd fragments (Chamov et al., 1999). The antibodies may be of human or non-human origin. Humanised and chimeric antibodies are also possible.

Fab fragments (fragment antigen binding = Fab) consist of the variable regions of both chains which are held together by the adjacent constant regions. They may be produced for example from conventional antibodies by treating with a protease such as papain or by DNA cloning. Other antibody fragments are $F(ab')_2$ fragments which can be produced by proteolytic digestion with pepsin.

By gene cloning it is also possible to prepare shortened antibody fragments which consist only of the variable regions of the heavy (VH) and light chain (VL). These are known as Fv fragments (fragment variable = fragment of the variable part). As covalent binding via the cystein groups of the constant chains is not possible in these Fv fragments, they are often stabilised by some other method. For this purpose the variable region of the heavy and light chains are often joined together by means of a short peptide fragment of about 10 to 30 amino acids, preferably 15 amino acids. This produces a single polypeptide chain in which VH and VL are joined together by a peptide linker. Such antibody fragments are also referred to as single chain Fv fragments (scFv). Examples of scFv antibodies are known and described, cf. for example Huston et al., 1988.

In past years various strategies have been developed for producing multimeric scFv derivatives. The intention is to produce recombinant antibodies with improved pharmacokinetic properties and increased binding avidity. In order to achieve the multimerisation of the scFv fragments they are produced as fusion proteins with multimerisation domains. The multimerisation domains may be, for example, the CH3 region of an IgG or helix structures ("coiled coil structures") such as the Leucine Zipper domains. In other strategies the interactions between the VH and VL regions of the scFv fragment are used for multimerisation (e.g. dia-, tri- and pentabodies).

The term diabody is used in the art to denote a bivalent homodimeric scFv derivative. Shortening the peptide linker in the scFv molecule to 5 to 10 amino acids results in the formation of homodimers by superimposing VH/VL chains. The diabodies may additionally be stabilised by inserted disulphite

bridges. Examples of diabodies can be found in the literature, e.g. in Perisic et al., 1994.

The term minibody is used in the art to denote a bivalent homodimeric scFv derivative. It consists of a fusion protein which contains the CH3 region of an immunoglobulin, preferably IgG, most preferably IgG1, as dimerisation region. This connects the scFv fragments by means of a hinge region, also of IgG, and a linker region. Examples of such minibodies are described by Hu et al., 1996.

The term triabody is used in the art to denote a trivalent homotrimeric scFv derivative (Kortt et al., 1997). The direct fusion of VH VL without the use of a linker sequence leads to the formation of trimers.

The fragments known in the art as mini antibodies which have a bi-, tri- or tetravalent structure are also derivatives of scFv fragments. The multimerisation is achieved by means of di-, tri- or tetrameric coiled coil structures (Pack et al., 1993 and 1995; Lovejoy et al., 1993).

Gene which codes for a fluorescent protein

The expression vector according to the invention contains a gene coding for a fluorescent protein functionally linked to the gene of interest and under the control of the hamster-ubiquitin/S27a promoter, a modified hamster-ubiquitin/S27a promoter or a homologue thereof.

The fluorescent protein may be, for example, a green, bluish-green, blue, yellow or other coloured fluorescent protein. One particular example is green fluorescent protein (GFP) obtained from *Aequorea victoria* or *Renilla reniformis* and mutants developed from them; cf. for example Bennet et al., 1998; Chalfie et al., 1994; WO 01/04306 and the literature cited therein.

Other fluorescent proteins and genes coding for them are described in WO 00/34318, WO 00/34326, WO 00/34526 and WO 01/27150 which are

incorporated herein by reference. These fluorescent proteins are fluorophores of non-bioluminescent organisms of the species *Anthozoa*, for example *Anemonia majano*, *Clavularia* sp., *Zoanthus* sp. I, *Zoanthus* sp. II, *Discosoma striata*, *Discosoma* sp. "red", *Discosoma* sp. "green", *Discosoma* sp. "Magenta", *Anemonia sulcata*.

The fluorescent proteins used according to the invention contain in addition to the wild-type proteins natural or genetically engineered mutants and variants, fragments, derivatives or variants thereof which have for example been fused with other proteins or peptides. The mutations used may for example alter the excitation or emission spectrum, the formation of chromophores, the extinction coefficient or the stability of the protein. Moreover, the expression in mammalian cells or other species can be improved by codon optimisation. According to the invention the fluorescent protein may also be used in fusion with a selectable marker, preferably an amplifiable selectable marker such as dihydrofolate reductase (DHFR).

The fluorescence emitted by the fluorescent proteins makes it possible to detect the proteins, e.g. by throughflow cytometry with a fluorescence-activated cell sorter (FACS) or by fluorescence microscopy.

Other regulatory elements

The hamster-ubiquitin/S27a promoter may be functionally combined with other regulatory sequences in order to increase/regulate the transcription activity in an expression cassette.

For example, the promoter may be functionally linked to enhancer sequences in order to increase the transcriptional activity. For this, one or more enhancers and/or several copies of an enhancer sequence may be used, e.g. a CMV or SV40 enhancer.

The term enhancer denotes a polynucleotide sequence which in the *cis* location acts on the activity of a promoter and thus stimulates the transcription

of a gene functionally connected to this promoter. Unlike promoters the effect of enhancers is independent of position and orientation and they can therefore be positioned in front of or behind a transcription unit, within an intron or even within the coding region. The enhancer may be located both in the immediate vicinity of the transcription unit and at a considerable distance from the promoter. It is also possible to have a physical and functional overlap with the promoter. The skilled man will be aware of a number of enhancers from various sources (and deposited in databanks such as GenBank, e.g. SV40 enhancers, CMV enhancers, polyoma enhancers, adenovirus enhancers) which are available as independent elements or elements cloned within polynucleotide sequences (e.g. deposited at the ATCC or from commercial and individual sources). A number of promoter sequences also contain enhancer sequences such as the frequently used CMV promoter. The human CMV enhancer is one of the strongest enhancers identified hitherto. One example of an inducible enhancer is the metallothionein enhancer, which can be stimulated by glucocorticoids or heavy metals.

Another possible modification is, for example, the introduction of multiple Sp1 binding sites. The promoter sequences may also be combined with regulatory sequences which allow control/regulation of the transcription activity. Thus, the promoter can be made repressible/ inducible. This can be done for example by linking to sequences which are binding sites for up- or down-regulating transcription factors. The above mentioned transcription factor Sp1, for example, has a positive effect on the transcription activity. Another example is the binding site for the activator protein AP1, which may act both positively and negatively on transcription. The activity of AP1 can be controlled by all kinds of factors such as, for example, growth factors, cytokines and serum (Faisst et al., 1992 and references therein). The transcription efficiency can also be increased by changing the promoter sequence by the mutation (substitution, insertion or deletion) of one, two, three or more bases and then determining, in a reporter gene assay, whether this has increased the promoter activity.

Basically, the additional regulatory elements include promoters other than the hamster-ubiquitin/S27a promoter, enhancers, termination and polyadenylation signals and other expression control elements. Both inducible and constitutively regulatory sequences are known for the various cell types. "Transcription-regulatory elements" generally comprise a promoter upstream of the gene sequence to be expressed, transcription initiation and termination sites and a polyadenylation signal.

The term promoter denotes a polynucleotide sequence which allows and controls the transcription of the genes or sequences functionally connected therewith. A promoter contains recognition sequences for binding RNA polymerase and the initiation site for transcription (transcription initiation site). In order to express a desired sequence in a certain cell type or a host cell a suitable functional promoter must be chosen. The skilled man will be familiar with a variety of promoters from various sources, including constitutive, inducible and repressible promoters. They are deposited in databanks such as GenBank, for example, and may be obtained as separate elements or elements cloned within polynucleotide sequences from commercial or individual sources. In inducible promoters the activity of the promoter may be reduced or increased in response to a signal. One example of an inducible promoter is the tetracycline (tet) promoter. This contains tetracycline operator sequences (tetO) which can be induced by a tetracycline-regulated transactivator protein (tTA). In the presence of tetracycline the binding of tTA to tetO is inhibited. Examples of other inducible promoters are the jun, fos, metallothionein and heat shock promoter (see also Sambrook et al., 1989; Gossen et al., 1994). Of the promoters which are particularly suitable for high expression in eukaryotes, there are for example the SV40 early promoter, adenovirus major late promoter, mouse metallothionein-I promoter, the long terminal repeat region of Rous Sarcoma Virus and the early promoter of human Cytomegalovirus. Examples of other heterologous mammalian promoters are the actin, immunoglobulin or heat shock promoter(s).

The term "transcription initiation site" refers to a nucleic acid in the construct which corresponds to the first nucleic acid which is incorporated in the primary

transcript, i.e. the mRNA precursor. The transcription initiation site may overlap with the promoter sequences.

The term "transcription termination site" refers to a nucleotide sequence which is normally at the 3' end of the gene of interest or of the gene section which is to be transcribed, and which brings about the termination of transcription by RNA polymerase.

The "polyadenylation signal" is a signal sequence which causes cleavage at a specific site at the 3' end of the eukaryotic mRNA and post-transcriptional incorporation of a sequence of about 100-200 adenine nucleotides (polyA tail) at the cleaved 3' end. The polyadenylation signal comprises the sequence AATAAA about 10-30 nucleotides upstream of the cleavage site and a sequence located downstream. Various polyadenylation elements are known such as tk polyA, SV40 late and early polyA or BGH polyA (described for example in US 5,122,458).

In a preferred embodiment of the present invention each transcription unit has a promoter or a promoter/enhancer element, a gene of interest and/or a marker gene as well as a transcription termination element. In another preferred embodiment the transcription unit contains two further translation regulatory units.

"Translation regulatory elements" comprise a translation initiation site (AUG), a stop codon and a polyA signal for each polypeptide to be expressed. For optimum expression it may be advisable to remove, add or change 5'- and/or 3'-untranslated regions of the nucleic acid sequence which is to be expressed, in order to eliminate any potentially unsuitable additional translation initiation codons or other sequences which might affect expression at the transcription or expression level. In order to promote expression, ribosomal consensus binding sites may alternatively be inserted immediately upstream of the start codon. In order to produce a secreted polypeptide the gene of interest usually contains a signal sequence which codes for a signal precursor peptide which transports the synthesised polypeptide to and through the ER membrane.

The signal sequence is often but not always located at the amino terminus of the secreted protein and is cleaved by signal peptidases after the protein has been inserted through the ER membrane. The gene sequence will usually but not necessarily contain its own signal sequence. If the native signal sequence is not present a heterologous signal sequence may be introduced in known manner. Numerous signal sequences of this kind are known to the skilled man and deposited in sequence databanks such as GenBank and EMBL.

One important regulatory element according to the invention is the internal ribosomal entry site (IRES). The IRES element comprises a sequence which functionally activates the translation initiation independently of a 5'-terminal methylguanosinium cap (CAP structure) upstream of the gene and in an animal cell allows the translation of two cistrons (open reading frames) from a single transcript. The IRES element provides an independent ribosomal entry site for the translation of the open reading frame located immediately downstream. In contrast to bacterial mRNA which may be multicistronic, i.e. it may code for numerous different polypeptides or products which are translated one after the other by the mRNA, the majority of mRNAs from animal cells are monocistronic and code for only one protein or product. In the case of a multicistronic transcript in a eukaryotic cell the translation would be initiated from the translation initiation site which was closest upstream and would be stopped by the first stop codon, after which the transcript would be released from the ribosome. Thus, only the first polypeptide or product coded by the mRNA would be produced during translation. By contrast, a multicistronic transcript with an IRES element which is functionally linked to the second or subsequent open reading frame in the transcript allows subsequent translation of the open reading frame located downstream thereof, so that two or more polypeptides or products coded by the same transcript are produced in the eukaryotic cell.

The IRES element may be of various lengths and various origins and may originate, for example, from the encephalomyocarditis virus (EMCV) or other Picorna viruses. Various IRES sequences and their use in the construction of vectors are described in the literature, cf. for example Pelletier et al., 1988;

Jang et al., 1989; Davies et al., 1992; Adam et al., 1991; Morgan et al., 1992; Sugimoto et al., 1994; Ramesh et al., 1996; Mosser et al., 1997.

The gene sequence located downstream is functionally linked to the 3' end of the IRES element, i.e. the spacing is selected so that the expression of the gene is unaffected or only marginally affected or has sufficient expression for the intended purpose. The optimum permissible distance between the IRES element and the start codon of the gene located downstream thereof for sufficient expression can be determined by simple experiments by varying the spacing and determining the expression rate as a function of the spacing using reporter gene assays.

By the measures described it is possible to obtain an optimum expression cassette which is of great value for the expression of heterologous gene products. An expression cassette obtained by means of one or more such measures is therefore a further subject of the invention.

Amplifiable Selectable Marker Gene

A preferred vector according to the invention additionally contains an amplifiable selectable marker gene which allows amplification of the amplifiable marker gene and preferably co-amplification of a transcription unit consisting of the hamster-ubiquitin/S27a gene, the gene of interest and the gene for the fluorescent protein. For this, the host cells transfected with a corresponding expression vector are cultivated in the presence of a suitable selection agent, so that only those host cells which have a number of gene copies of at least the amplifiable selectable marker gene can replicate (be replicated). Preferably, this is achieved by stepwise cultivation of the cells in the presence of increasing amounts of selecting agent.

The amplifiable selectable marker gene usually codes for an enzyme which is needed for the growth of eukaryotic cells under certain cultivation conditions. For example, the amplifiable selectable marker gene may code for dihydrofolate reductase (DHFR). In this case the gene is amplified if a host

cell transfected therewith is cultivated in the presence of the selecting agent methotrexate (MTX).

The following Table 1 gives examples of other amplifiable selectable marker genes and the associated selecting agents which may be used according to the invention, which are described in an overview by Kaufman, Methods in Enzymology, 185:537-566 (1990).

Table 1: Amplifiable selectable marker genes

Amplifiable selectable marker gene	Accession number	Selecting agent
dihydrofolate reductase	M19869 (hamster) E00236 (mouse)	methotrexate (MTX)
metallothionein	D10551 (hamster) M13003 (human) M11794 (rat)	cadmium
CAD (carbamoylphosphate synthetase : aspartate transcarbamylase: dihydroorotase)	M23652 (hamster) D78586 (human)	N-phosphoacetyl-L-aspartate
adenosine-deaminase	K02567 (human) M10319 (mouse)	Xyl-A- or adenosine, 2'-deoxycytosine
AMP (adenylate)-deaminase	D12775 (human) J02811 (rat)	adenine, azaserine, coformycin
UMP-synthase	J03626 (human)	6-azauridine, pyrazofuran
IMP 5'-dehydrogenase	J04209 (hamster) J04208 (human) M33934 (mouse)	mycophenolic acid
xanthine-guanine-phosphoribosyltransferase	X00221 (E. coli)	mycophenolic acid with limiting xanthine
mutant HGPRTase or mutant thymidine-kinase	J00060 (hamster) M13542, K02581 (human) J00423, M68489 (mouse) M63983 (rat) M36160 (Herpes virus)	hypoxanthine, aminopterin and thymidine (HAT)
thymidylate-synthetase	D00596 (human) M13019 (mouse) L12138 (rat)	5-fluorodeoxyuridine
P-glycoprotein 170 (MDR1)	AF016535 (human) J03398 (mouse)	several drugs, e.g. adriamycin, vincristine, colchicine
ribonucleotide reductase	M124223, K02927 (mouse)	aphidicoline

glutamine-synthetase	AF150961 (hamster) U09114, M60803 (mouse) M29579 (rat)	methionine sulphoximine (MSX)
asparagine-synthetase	M27838 (hamster) M27396 (human) U38940 (mouse) U07202 (rat)	β -aspartylhydroxamate, albizzin, 5'azacytidine
argininosuccinate- synthetase	X01630 (human) M31690 (mouse) M26198 (bovine)	canavanin
ornithine-decarboxylase	M34158 (human) J03733 (mouse) M16982 (rat)	α -difluoromethylornithine
HMG-CoA-reductase	L00183, M12705 (hamster) M11058 (human)	compactin
N-acetylglucosaminyl- transferase	M55621 (human)	tunicamycin
threonyl-tRNA-synthetase	M63180 (human)	borrelidin
Na ⁺ K ⁺ -ATPase	J05096 (human) M14511 (rat)	ouabain

According to the invention the amplifiable selectable marker gene used is preferably a gene which codes for a polypeptide with the function of DHFR, e.g. for DHFR or a fusion protein from the fluorescent protein and DHFR.

DHFR is necessary for the biosynthesis of purines. Cells which lack the DHFR genes cannot grow in purine-deficient medium. The DHFR gene is therefore a useful selectable marker for selecting and amplifying genes in cells cultivated in purine-free medium. The selecting medium used in conjunction with the DHFR gene is methotrexate (MTX). The present invention therefore includes a method of preparing highly productive recombinant host cells which contains the following steps: (i) transfection of the host cells with genes which code at least for a protein of interest, a fluorescent protein and DHFR; (ii) cultivation of the cells under conditions which allow expression of the various genes, and (iii) the amplification of the co-integrated genes by cultivating the cells in the presence of a selecting agent which allows the amplification of at least the amplifiable selectable marker gene such as methotrexate. Preferably the transfected cells are cultivated in hypoxanthine/thymidine-free medium in the absence of serum and with the addition of increasing concentrations of MTX. Preferably the

concentration of MTX in the first amplification step is at least 200 nM and in a more preferred embodiment it is at least 500 nM and may be increased step by step to 1 μ M. In individual cases concentrations of more than 1 μ M may be used.

Mammalian cells, preferably mouse myeloma and hamster cells, are preferred host cells for the use of DHFR as an amplifiable selectable marker. The cell lines CHO-DUKX (ATCC CRL-9096) and CHO-GD44 (Urlaub et al., 1983) are particularly preferred as they have no DHFR activity of their own, as a result of mutation. In order to be able to use the DHFR-induced amplification in other cell types as well which have their own endogenous DHFR activity, it is possible to use in the transfection process a mutated DHFR gene which codes for a protein with reduced sensitivity to methotrexate (Simonson et al., 1983; Wigler et al., 1980; Haber et al., 1982).

Preparation of expression vectors according to the invention

The expression vector according to the invention may theoretically be prepared by conventional methods known in the art, as described by Sambrook et al. (1989), for example. Sambrook also describes the functional components of a vector, e.g. suitable promoters (in addition to the hamster ubiquitin/S27a promoter), enhancers, termination and polyadenylation signals, antibiotic resistance genes, selectable markers, replication starting points and splicing signals. Conventional cloning vectors may be used to produce them, e.g. plasmids, bacteriophages, phagemids, cosmids or viral vectors such as baculovirus, retroviruses, adenoviruses, adeno-associated viruses and herpes simplex virus, as well as artificial chromosomes/mini chromosomes. The eukaryotic expression vectors typically also contain prokaryotic sequences such as, for example, replication origin and antibiotic resistance genes which allow replication and selection of the vector in bacteria. A number of eukaryotic expression vectors which contain multiple cloning sites for the introduction of a polynucleotide sequence are known and some may be obtained commercially from various companies such as Stratagene, La Jolla,

CA, USA; Invitrogen, Carlsbad, CA, USA; Promega, Madison, WI, USA or BD Biosciences Clontech, Palo Alto, CA, USA.

The hamster-ubiquitin/S27a promoter, the gene of interest, the gene coding for a fluorescent protein and preferably also the amplifiable selectable marker gene, e.g. dihydrofolate reductase, and optionally additional regulatory elements such as the internal ribosomal entry site (IRES), enhancers or a polyadenylation signal are introduced into the expression vector in a manner familiar to those skilled in the art. An expression vector according to the invention contains, at the minimum, a ubiquitin/S27a promoter, the gene of interest and the gene coding for a fluorescent protein. Preferably, the expression vector also contains an amplifiable selectable marker gene. According to the invention, modified ubiquitin/S27a promoters, e.g. the modified ubiquitin/S27a promoters described in the present application, are also used. Particularly preferred is an expression vector in which the ubiquitin promoter, the gene of interest and the gene which codes for a fluorescent protein are functionally linked together or are functionally linked.

Within the scope of the present description the term "functional linking" or "functionally linked" refers to two or more nucleic acid sequences or partial sequences which are positioned so that they can perform their intended function. For example, a promoter/enhancer is functionally linked to a coding gene sequence if it is able to control or modulate the transcription of the linked gene sequence in the *cis* position. Generally, but not necessarily, functionally linked DNA sequences are close together and, if two coding gene sequences are linked or in the case of a secretion signal sequence, in the same reading frame. Although a functionally linked promoter is generally located upstream of the coding gene sequence it does not necessarily have to be close to it. Enhancers need not be close by either, provided that they assist the transcription of the gene sequence. For this purpose they may be both upstream and downstream of the gene sequence, possibly at some distance from it. A polyadenylation site is functionally linked to a gene sequence if it is positioned at the 3' end of the gene sequence in such a way that the transcription progresses via the coding sequence to the polyadenylation

signal. Linking may take place according to conventional recombinant methods, e.g. by the PCR technique, by ligation at suitable restriction cutting sites or by splicing. If no suitable restriction cutting sites are available synthetic oligonucleotide linkers or adaptors *per se* may be used in a manner known. According to the invention the functional linking preferably does not take place via intron sequences.

In one of the embodiments described, the ubiquitin/S27a promoter or a modified form thereof, the gene of interest and the gene coding for a fluorescent protein are functionally linked together. This means for example that both the gene of interest and the gene coding for a fluorescent protein are expressed starting from the same ubiquitin/S27a promoter or a modified form thereof. In a particularly preferred embodiment the functional linking takes place via an IRES element, so that a bicistronic mRNA is synthesised from both genes. The expression vector according to the invention may additionally contain enhancer elements which act functionally on one or more promoters. Particularly preferred is an expression vector in which the ubiquitin/S27a promoter or a modified form thereof is linked to an enhancer element, e.g. an SV40 enhancer or a CMV enhancer element.

Fundamentally, the expression of the genes within an expression vector may take place starting from one or more transcription units. The term transcription unit is defined as a region which contains one or more genes to be transcribed. The genes within a transcription unit are functionally linked to one another in such a way that all the genes within such a unit are under the transcriptional control of the same promoter or promoter/ enhancer. As a result of this transcriptional linking of genes, more than one protein or product can be transcribed from a transcription unit and thus expressed. Each transcription unit contains the regulatory elements which are necessary for the transcription and translation of the gene sequences contained therein. Each transcription unit may contain the same or different regulatory elements. IRES elements or introns may be used for the functional linking of the genes within a transcription unit.

The expression vector may contain a single transcription unit for expressing the gene of interest, the gene for the fluorescent protein and the amplifiable selectable marker. Alternatively, these genes may also be arranged in two or more transcription units. Various combinations of the genes within a transcription unit are possible. In another embodiment of the present invention more than one expression vector consisting of one, two or more transcription units may be inserted in a host cell by cotransfection or in successive transfections in any desired order. Any combination of regulatory elements and genes on each vector can be selected provided that adequate expression of the transcription units is ensured. If necessary, other regulatory elements and genes, e.g. additional genes of interest or selectable markers, may be positioned on the expression vectors.

Accordingly, the expression vector according to the invention may contain the gene which codes for a fluorescent protein and the amplifiable selectable marker gene in one or in two separate transcription units. Each transcription unit can transcribe and express one or more gene products. If both genes are contained in one transcription unit they are under the control of the same promoter or promoter/ enhancer, while preferably an IRES element is used to ensure the functional linking of all the components. If the gene which codes for a fluorescent protein and the amplifiable selectable marker gene are contained in two separate transcription units, they may be under the control of the same or different promoters/enhancers. However, preferably, its natural or a weaker heterologous promoter, e.g. SV40 early promoter, is used for the selectable marker gene and preferably no enhancer is used. Expression vectors with two separate transcription units are preferred within the scope of the invention. One (bicistronic) transcription unit contains the gene of interest and the gene coding for a fluorescent protein, while the other transcription unit contains the amplifiable selectable marker gene. Preferably, each transcription unit is limited at the 3' end by a sequence which codes for a polyA signal, preferably tk polyA, BGH polyA or SV40 polyA.

Also preferred according to the invention are those vectors which instead of the gene of interest have only a multiple cloning site which allows the cloning

of the gene of interest via recognition sequences for restriction endonucleases. Numerous recognition sequences for all kinds of restriction endonucleases as well as the associated restriction endonucleases are known from the prior art. Preferably, sequences are used which consist of at least six nucleotides as recognition sequence. A list of suitable recognition sequences can be found for example in Sambrook et al., 1989.

Host Cells

For transfection with the expression vector according to the invention eukaryotic host cells are used, preferably mammalian cells and more particularly rodent cells such as mouse, rat and hamster cell lines. The successful transfection of the corresponding cells with an expression vector according to the invention results in transformed, genetically modified, recombinant or transgenic cells, which are also the subject of the present invention.

Preferred host cells for the purposes of the invention are hamster cells such as BHK21, BHK TK⁻, CHO, CHO-K1, CHO-DUKX, CHO-DUKX B1 and CHO-DG44 cells or derivatives/descendants of these cell lines. Particularly preferred are CHO-DG44, CHO-DUKX, CHO-K1 and BHK21 cells, particularly CHO-DG44 and CHO-DUKX cells. Also suitable are myeloma cells from the mouse, preferably NS0 and Sp2/0 cells and derivatives/descendants of these cell lines.

Examples of hamster and mouse cells which can be used according to the invention are given in Table 2 that follows. However, derivatives and descendants of these cells, other mammalian cells including but not restricted to cell lines of humans, mice, rats, monkeys, rodents, or eukaryotic cells, including but not restricted to yeast, insect and plant cells, may also be used as host cells for the production of biopharmaceutical proteins.

Table 2: Hamster and Mouse Production Cell Lines

Cell line	Accession Number
NS0	ECASS No. 85110503
Sp2/0-Ag14	ATCC CRL-1581
BHK21	ATCC CCL-10
BHK TK ⁻	ECACC No. 85011423
HaK	ATCC CCL-15
2254-62.2 (BHK-21-derivative)	ATCC CRL-8544
CHO	ECACC No. 8505302
CHO-K1	ATCC CCL-61
CHO-DUKX (= CHO duk ⁻ , CHO/dhfr ⁻)	ATCC CRL-9096
CHO-DUKX B1	ATCC CRL-9010
CHO-DG44	Urlaub et al; Cell 32[2], 405-412, 1983
CHO Pro-5	ATCC CRL-1781
V79	ATCC CCC-93
B14AF28-G3	ATCC CCL-14
CHL	ECACC No. 87111906

The transfection of the eukaryotic host cells with a polynucleotide or one of the expression vectors according to the invention is carried out by conventional methods (Sambrook et al., 1989; Ausubel et al., 1994). Suitable methods of transfection include for example liposome-mediated transfection, calcium phosphate coprecipitation, electroporation, polycation- (e.g. DEAE dextran)-mediated transfection, protoplast fusion, microinjection and viral infections. According to the invention stable transfection is preferably carried out in which the constructs are either integrated into the genome of the host cell or an artificial chromosome/minichromosome, or are episomally contained in stable manner in the host cell. The transfection method which gives the

optimum transfection frequency and expression of the heterologous gene in the host cell in question is preferred. By definition, every sequence or every gene inserted in a host cell is referred to as a "heterologous sequence" or "heterologous gene" in relation to the host cell. This applies even if the sequence to be introduced or the gene to be introduced is identical to an endogenous sequence or an endogenous gene of the host cell. For example, a hamster actin gene introduced into a hamster host cell is by definition a heterologous gene.

In the recombinant production of heterodimeric proteins such as e.g. monoclonal antibodies (mAb), the transfection of suitable host cells can theoretical be carried out by two different methods. mAb's of this kind are composed of a number of subunits, the heavy and light chains. Genes coding for these subunits may be accommodated in independent or multicistronic transcription units on a single plasmid with which the host cell is then transfected. This is intended to secure the stoichiometric representation of the genes after integration into the genome of the host cell. However, in the case of independent transcriptional units it must hereby be ensured that the mRNAs which encode the different proteins display the same stability and transcriptional and translational efficiency. In the second case, the expression of the genes take place within a multicistronic transcription unit by means of a single promoter and only one transcript is formed. By using IRES elements, a highly efficient internal translation initiation of the genes is obtained in the second and subsequent cistrons. However, the expression rates for these cistrons are lower than that of the first cistron, the translation initiation of which, by means of a so-called "cap"-dependent pre-initiation complex, is substantially more efficient than IRES-dependent translation initiation. In order to achieve a truly equimolar expression of the cistrons, additional inter-cistronic elements may be introduced, for example, which ensure uniform expression rates in conjunction with the IRES elements (WO 94/05785).

Another possible way of simultaneously producing a number of heterologous proteins, which is preferred according to the invention, is cotransfection, in

which the genes are separately integrated in different expression vectors. This has the advantage that certain proportions of genes and gene products with one another can be adjusted, thereby balancing out any differences in the mRNA stability and in the efficiency of transcription and translation. In addition, the expression vectors are more stable because of their small size and are easier to handle both during cloning and during transfection.

In one particular embodiment of the invention, therefore, the host cells are additionally transfected, preferably cotransfected, with one or more vectors having genes which code for one or more other proteins of interest. The other vector or vectors used for the cotransfection code, for example, for the other protein or proteins of interest under the control of the same promoter/enhancer combination and for at least one other selectable marker, for example neomycin phosphotransferase.

According to the invention the host cells are preferably established, adapted and cultivated under serum-free conditions, optionally in media which are free from animal proteins/peptides. Examples of commercially obtainable media include Ham's F12 (Sigma, Deisenhofen, DE), RPMI-1640 (Sigma), Dulbecco's Modified Eagle's Medium (DMEM; Sigma), Minimal Essential Medium (MEM; Sigma), Iscove's Modified Dulbecco's Medium (IMDM; Sigma), CD-CHO (Invitrogen, Carlsbad, CA, USA), CHO-S-SFMII (Invitrogen), serum-free CHO-Medium (Sigma) and protein-free CHO-Medium (Sigma). Each of these media may optionally be supplemented with various compounds, e.g. hormones and/or other growth factors (e.g. insulin, transferrin, epidermal growth factor, insulin-like growth factor), salts (e.g. sodium chloride, calcium, magnesium, phosphate), buffers (e.g. HEPES), nucleosides (e.g. adenosine, thymidine), glutamine, glucose or other equivalent nutrients, antibiotics and/or trace elements. Although serum-free media are preferred according to the invention, the host cells may also be cultivated using media which have been mixed with a suitable amount of serum. In order to select genetically modified cells which express one or more selectable marker genes, one or more selecting agents are added to the medium.

The term "selecting agent" refers to a substance which affects the growth or survival of host cells with a deficiency for the selectable marker gene in question. For example, in order to select for the presence of an expressed antibiotic resistance gene such as neomycin phosphotransferase, the antibiotic geneticin (G418) is preferably used as the medium additive. The selecting agent may also be a substance which triggers amplification of the selectable marker gene if the gene used is an amplifiable selectable marker gene (cf. Table 1). Methotrexate, for example, is a selecting medium which is suitable for amplifying the DHFR gene. Examples of other selecting agents which trigger amplification are listed in Table 1.

A selectable marker gene is a gene which allows the specific selection of cells which contain this gene by the addition of a corresponding selecting agent to the cultivation medium. As an illustration, an antibiotic resistance gene may be used as a positive selectable marker. Only cells which have been transformed with this gene are able to grow in the presence of the corresponding antibiotic and are thus selected. Untransformed cells, on the other hand, are unable to grow or survive under these selection conditions. There are positive, negative and bifunctional selectable markers. Positive selectable markers permit the selection and hence enrichment of transformed cells by conferring resistance to the selecting agent or by compensating for a metabolic or catabolic defect in the host cell. By contrast, cells which have received the gene for the selectable marker can be selectively eliminated by negative selectable markers. An example of this is the thymidine kinase gene of the Herpes Simplex virus, the expression of which in cells with the simultaneous addition of acyclovir or gancyclovir leads to the elimination thereof. The selectable markers used in this invention, including the amplifiable selectable markers, include genetically modified mutants and variants, fragments, functional equivalence, derivatives, homologues and fusions with other proteins or peptides, provided that the selectable marker retains its selective qualities. Such derivatives display considerable homology in the amino acid sequence in the regions or domains which are deemed to be selective. The literature describes a large number of selectable marker

genes including bifunctional (positive/negative) markers (see for example WO 92/08796 and WO 94/28143). Examples of selectable markers which are usually used in eukaryotic cells include the genes for aminoglycoside phosphotransferase (APH), hygromycin phosphotransferase (HYG), dihydrofolate reductase (DHFR), thymidine kinase (TK), glutamine synthetase, asparagine synthetase and genes which confer resistance to neomycin (G418), puromycin, histidinol D, belomycin, phleomycin and zeocin.

It is also possible to select transformed cells by fluorescence-activated cell sorting (FACS). For this, bacterial β -galactosidase, cell surface markers or fluorescent proteins may be used (e.g. green fluorescent protein (GFP) and the variants thereof from *Aequorea victoria* and *Renilla reniformis* or other species, red fluorescent protein and proteins which fluoresce in other colours and their variants from non-bioluminescent organisms such as e.g. *Discosoma sp.*, *Anemonia sp.*, *Clavularia sp.*, *Zoanthus sp.*) for the selection of transformed cells.

In the present invention, the use of the DHFR gene is preferred for the selection of genetically modified (recombinant) host cells as the amplifiable selectable marker gene. This marker is particularly suitable for the selection and subsequent amplification when using DHFR negative basic cells such as CHO-DG44 or CHO-DUKX, as these cells do not express endogenous DHFR and therefore do not grow in purine-free medium. Consequently, the DHFR gene may be used here as a dominant selectable marker and the transformed cells are selected in hypoxanthine/ thymidine-free medium. In order to achieve DHFR-mediated gene amplification methotrexate (MTX) is used. The growth properties are critically influenced by the addition of MTX. Usually, a substantial deterioration is observed in the fermentation robustness of the cells as the MTX concentration and amplification step increase. Surprisingly, however, it has been found that, using the clone selection system according to the invention, recombinant host cells can be enriched which display much more robust behaviour in the presence of high concentrations of MTX (cf Fig. 7). Thus, host cells which have been identified and sorted using a Fluorescence-Activated Cell Sorter (FACS) were able to be cultivated and

amplified in the presence of 500 nM, preferably in the presence of 1 μ M MTX, which resulted in a significant increase in productivity. Thus, a method of selecting highly productive host cells is regarded as being particularly according to the invention if host cells which are transfected with an expression vector according to the invention and express at least the gene of interest, the fluorescent protein and a DHFR gene are sorted by FACS sorting, and are subjected to a gene amplification step in the presence of at least 500 nM, preferably 1 μ M of MTX.

By "fermentation robustness" is meant the growth properties of the cells, such as for example the maintenance of certain growth rates, robustness to "upscaling" (larger sizes of bioreactors) and the achievement of high cell counts and vitalities in the maintenance of the stock in order to meet the industrial production rates on upscaling.

Expression

The term expression relates to the transcription and/or translation of a heterologous gene sequence in a host cell. The expression rate can be generally determined, either on the basis of the quantity of corresponding mRNA which is present in the host cell or on the basis of the quantity of gene product produced which is encoded by the gene of interest. The quantity of mRNA produced by transcription of a selected nucleotide sequence can be determined for example by northern blot hybridisation, ribonuclease-RNA-protection, *in situ* hybridisation of cellular RNA or by PCR methods (Sambrook et al., 1989; Ausubel et al., 1994). Proteins which are encoded by a selected nucleotide sequence can also be determined by various methods such as, for example, ELISA, western blot, radioimmunoassay, immunoprecipitation, detection of the biological activity of the protein or by immune staining of the protein followed by FACS analysis (Sambrook et al., 1989; Ausubel et al., 1994).

The terms "high expression level (or rate), high expression, increased expression or high productivity" refer to the long-lasting and sufficiently high

expression or synthesis of a heterologous sequence introduced into a host cell, e.g. of a gene coding for a therapeutic protein. Increased or high expression or a high expression level or rate or a high productivity are present if a cell according to the invention is cultivated by one of the methods according to the invention described here and if this cell produces at least more than roughly 5 pg of the desired gene product per day (5 pg/day/cell). Increased or high expression or a high expression level or rate or a high productivity are also present if the cell according to the invention produces at least more than roughly 10 pg of the desired gene product per day (10 pg/day/cell). Increased or high expression or a high expression level or rate or high productivity are present in particular if the cell according to the invention produces at least more than roughly 15 pg of the desired gene product per day (15 pg/day/cell). Increased or high expression or a high expression level or rate or high productivity are present in particular if the cell according to the invention produces at least more than roughly 20 pg of the desired gene product per day (20 pg/day/cell). Particularly increased or high expression or a particularly high expression level or rate or particularly high productivity are present if the cell according to the invention produces at least more than roughly 30 pg of the desired gene product per day (30 pg/day/cell).

High or increased expression, high productivity or a high expression level or rate according to the invention can be achieved in various ways. For example, through coexpression of the gene of interest with a gene for an amplifiable selectable marker it is possible to select and identify cells which express the heterologous gene to a high degree. The amplifiable selectable marker not only allows the selection of stably transfected host cells but also the gene amplification of the heterologous gene of interest. The additional copies of the nucleic acids may be integrated into the genome of the host cells, into additional artificial/mini-chromosomes or into episomally located polynucleotides. This procedure may be combined with a FACS-assisted selection of recombinant host cells which contain, as additional selectable marker, one or more fluorescent proteins (e.g. GFP) or a cell surface marker. Other methods of obtaining increased expression, and a combination of different methods may also be used, are based for example on the use of

(artificial) transcription factors, treatment of the cells with natural or synthetic agents for up-regulating endogenous or heterologous gene expression, improving the stability (half-life) of mRNA or the protein, improving the initiation of mRNA translation, increasing the gene dose by the use of episomal plasmids (based on the use of viral sequences as replication origins, e.g. SV40, polyoma, adenovirus, EBV or BPV), the use of amplification-promoting sequences (Hermann et al., 1994) or in vitro amplification systems based on DNA concatemers (Monaco et al., 1996).

According to the invention, coupled transcription of the gene of interest and of the gene which codes for the fluorescent protein is carried out. The resulting bicistronic mRNA expresses both the protein of interest and the fluorescent protein. On the basis of this coupling of the expression of the protein of interest and the fluorescent protein it is easily possible according to the invention to select and isolate high-producing recombinant host cells by means of the fluorescent protein expressed, e.g. by sorting using a fluorescence activated cell sorter (FACS).

The selection of recombinant host cells which exhibit high vitality and an increased expression rate of the desired gene product is a multistage process. The host cells which have been transfected with the expression vector according to the invention or optionally cotransfected with another vector, for example, are investigated at least for the expression of the gene which codes for a fluorescent protein and is coupled to the gene of interest, in order to identify and select the cells/cell population which exhibit the highest expression rates of fluorescent protein. Preferably, only the cells which belong to the 10% of cells with the highest expression rate of fluorescent protein are sorted out and further cultivated. In practice this means that the brightest 10% of the fluorescent cells are sorted out and further cultivated. Accordingly, the brightest 5%, preferably the brightest 3% or even the brightest 1% of the fluorescent cells of a cell mixture can also be sorted out and replicated. In a particularly preferred embodiment only the brightest 0.5% or the brightest 0.1% of the fluorescent cells are sorted out and replicated.

For this purpose, the cells which have previously been transformed with the expression vector according to the invention are cultivated in a selection medium which optionally also contains a selecting agent specific for the amplifiable selectable marker. Concentrations of selecting agent increasing step by step may be used to exert a gradually increasing selection pressure.

The selection step may be carried out on cell pools or using pre-sorted cell pools/cell clones. One or more, preferably two or more and especially three or more sorting steps may be carried out, while between the individual sorting steps the cells may be cultivated and replicated for a specific length of time, e.g. roughly two weeks in the case of pools.

If desired the host cells may be subjected to one or more gene amplification steps in order to increase the copy number of at least the gene of interest and the amplifiable selectable marker gene. Processes for step-by-step gene amplification using methotrexate are described for example in US 5,179,017. According to the invention the high productivity which can be achieved is not tied to a high number of gene copies. Rather, it is the expression of increased stability and fermentation robustness in the high-performance clones. It is therefore possible to reduce the number of gene amplification steps required and to carry out only a single gene amplification, for example.

Accordingly, the invention thus relates to a method of selecting cells which comprises the following steps:

- (i) transformation of suitable host cells at least with one of the vectors according to the invention, wherein the DNA of the expression vectors is preferably stably incorporated in the host cell genome or in artificial chromosomes/minichromosomes;
- (ii) the transformed cells are cultivated under conditions which allow expression of the gene of interest and of the fluorescent protein;
- (iii) the cells are cultivated in the presence of at least one selecting agent so that only those cells which are able to grow in the presence of the selecting agent are replicated;

- (iv) the cells which exhibit the highest expression rate of fluorescent protein are sorted from a cell mixture, the cells being detected and sorted by means of a fluorescence-activated cell sorter (FACS);
- (v) the sorted cells with the highest expression rates for the fluorescent protein are cultivated.

Optionally, steps ii) - v) can be repeated one or more times with the cells obtained according to step v). Moreover, the transformed cells may optionally also be subjected to one or more gene amplification steps by being cultivated in the presence of a selecting agent which leads to amplification of the amplifiable selectable marker gene. This step may be carried out both with cells which have not yet been sorted and also with cells which have already been pre-sorted one or more times.

The invention also relates to a process in which correspondingly sorted cells are replicated and used to prepare the coding gene product of interest. The selected high-producing cells are preferably cultivated in a serum-free culture medium and preferably in suspension culture under conditions which allow expression of the gene of interest. The protein of interest is preferably obtained from the cell culture medium as a secreted gene product. When the protein is expressed without a secretion signal, however, the gene product may also be isolated from cell lysates. In order to obtain a pure, homogeneous product which is substantially free from other recombinant proteins and host cell proteins, conventional purification steps are carried out. First of all, usually, cells and cell debris are removed from the culture medium or lysate. The desired gene product can then be freed from contaminating soluble proteins, polypeptides and nucleic acids, e.g. by fractionation on immunoaffinity and ion exchange columns, ethanol precipitation, reversed phase HPLC or chromatography on Sephadex, silica or cation exchange resins such as DEAE. Methods for purifying a heterologous protein expressed by recombinant cells are known to the skilled man and described in the literature, e.g. by Harris et al. (1995) and Scopes (1988).

The invention is hereinafter explained more fully with reference to some non-restrictive exemplifying embodiments.

EXAMPLES

Abbreviations

AP:	alkaline phosphatase
bp:	base pair
CHO:	Chinese hamster ovary
DHFR:	dihydrofolate-reductase
ELISA:	enzyme-linked immunosorbant assay
FACS:	fluorescence-activated cell sorter
FAP:	fibroblast-activated protein
GFP:	green fluorescent protein
HBSS:	Hanks Balanced Salt Solution
HT:	hypoxanthine/thymidine
HRPO:	horseradish peroxidase
IRES:	internal ribosomal entry site
kb:	Kilobase
mAb:	monoclonal antibody
MTX:	methotrexate
PCR:	polymerase chain reaction
sICAM:	soluble intracellular adhesion molecule

Methods

1. Cell culture and Transfection

The cells CHO-DG44/dhfr-/- (Urlaub et al., 1983) were permanently cultivated as suspension cells in serum-free CHO-S-SFMII medium supplemented with hypoxanthine and thymidine (Invitrogen GmbH, Karlsruhe, DE) in cell culture flasks at 37°C in a damp atmosphere and 5% CO₂. The cell counts and viability were determined with a CASY1 Cell Counter (Schaerfe System, DE) or by tryptan blue staining and the cells were then seeded in a concentration of 1 – 3 x10⁵/mL and run every 2 – 3 days.

Lipofectamine Plus Reagent (Invitrogen GmbH) was used for the transfection of CHO-DG44. For each transfection mixture a total of 1 µg of plasmid-DNA, 4 µL of lipofectamine and 6 µL of Plus reagent were mixed together according to the manufacturer's instructions and added in a volume of 200 µL to 6 x10⁵ exponentially growing CHO-DG44 cells in 0.8 mL of HT-supplemented CHO-S-SFMII medium. After three hours' incubation at 37°C in a cell incubator 2 mL of HT-supplemented CHO-S-SFMII medium was added. For the DHFR-based selection of stably transfected CHO-DG44 the cells were transferred 2 days after transfection into CHO-S-SFMII medium with no added hypoxanthine and thymidine, changing the medium every 3 to 4 days. In a DHFR- and neomycin phosphotransferase-based selection in the case of co-transfection in which one expression vector contained a DHFR and the other expression vector contained a neomycin-phosphotransferase selectable marker, G418 (Invitrogen) was also added to the medium in a concentration of 400µg/mL.

A DHFR-based gene amplification of the integrated heterologous genes was obtained by the addition of the selecting agent MTX (Sigma, Deisenhofen, DE) in a concentration of 5 – 2000 nM to the HT-free CHO-S-SFMII medium.

2. Expression vectors

To analyse the expression, eukaryotic expression vectors were used which are based on the pAD-CMV vector (Werner et al., 1998) and mediate the constitutive expression of a heterologous gene by the combination of CMV enhancer/hamster ubiquitin/S27a promoter (WO 97/15664). While the base vector pBID contains the DHFR minigene which acts as an amplifiable selectable marker (cf e.g. EP 0 393 438), in the vector pBIN the DHFR minigene has been replaced by a neomycin resistance gene (Fig.2). For this purpose the selectable marker neomycin-phosphotransferase, including SV40 early promoter and TK-polyadenylation signal, was isolated from the commercial plasmid pBK-CMV (Stratagene, La Jolla, CA, USA) as a 1640 bp Bsu36I fragment. After a reaction to fill in the ends of the fragment with Klenow-DNA-polymerase the fragment was ligated with the 3750 bp

Bsu36I/StuI fragment of the vector pBID, which was also treated with Klenow-DNA-polymerase.

In the bicistronic base vector pBIDG (Fig. 2) the IRES-GFP gene region was isolated from the vector pIRES2-EGFP (Clontech, Palo Alto, CA, USA) and brought under the control of the CMV enhancer/promoter in the vector pBID so that the multiple cloning site between the promoter region and IRES-element was retained. The following procedure was used. In a PCR mutagenesis in which the plasmid pIRES2-EGFP acted as the template, on the one hand the HindIII cutting site AAGCTT within the IRES sequence was converted into the sequence ATGCTT by the use of mutagenic primers and thus eliminated. On the other hand an XbaI cutting site was inserted by means of a primer with complementarity to the 5'end of the IRES sequence or a SphI cutting site was introduced by means of a primer with complementarity to the 3'end of the GFP sequence. The resulting PCR fragment, which contained the complete IRES and GFP sequence, was digested with XbaI and SphI and cloned into the singular XbaI cutting site at the 3'end of the multiple cloning site of the vector pBID.

The human sICAM gene was isolated as a HindIII/Sall fragment from pAD-sICAM (Werner et al., 1998) and cloned into the corresponding cutting sites of the vector pBIDG, resulting in the vector pBIDG-sICAM (Fig. 3).

In order to express the monoclonal humanised F19 antibody the heavy chain was isolated as a 1.5 kb Nael/HindIII fragment from the plasmid pG1D105F19HC (NAGENESEQ: AAZ32786) and cloned into the vector pBIDG digested with EcoRI (filled in with Klenow-DNA-polymerase) and HindIII, resulting in the vector pBIDG-F19HC (Fig.3). The light chain on the other hand was isolated as a 1.3 kb HindIII/EcoRI fragment from the plasmid pKN100F19LC (NAGENESEQ: AAZ32784) and cloned into the corresponding cutting sites of the vector pBIN, producing the vector pBIN-F19LC (Fig.3).

3. FACS

The flow-cytometric analyses and sorting were carried out with a Coulter Epics Altra device. The FACS is fitted with a helium-argon laser with an excitation wavelength of 488 nm. The fluorescence intensity is absorbed at a wavelength suited to the fluorescence protein and process by means of the attached software Coulter Expo32. The sorting was carried out at a rate of 8000 – 10000 events/second. The suspended cells were centrifuged (5 min at 180xg) and adjusted to a cell concentration of 1 – 1.5 x10⁷/mL in HBSS. Then the cells were sorted according to their fluorescence protein signal. The cells were taken up in test tubes already containing culture medium, then centrifuged and seeded into suitable culture vessels depending on the number of cells sorted.

4. *ELISA*

The sICAM titres in supernatants of stably transfected CHO-DG44 cells were quantified by ELISA according to standard procedures (Ausubel et al., 1994, updated), using on the one hand a goat anti human IgG Fc fragment (Dianova, Hamburg, DE) and on the other hand an AP-conjugated goat anti human kappa light chain antibody (Sigma). Purified F19 antibody was used as the standard.

Productivities (pg/cell/day) were calculated by the formula $\text{pg}/((\text{Ct}-\text{Co}) \text{ t} / \text{ln}(\text{Ct}-\text{Co}))$, where Co and Ct are the cell count on seeding and harvest, respectively, and t is the cultivation time.

Example 1: Comparison of the CMV and hamster-ubiquitin/S27a-promoter activity

In order to compare the activity of the hamster-ubiquitin/S27a-promoters with that of the CMV promoter frequently used in eukaryotic expression vectors, CHO-DG44 cells were transfected with various recombinant vectors. The heterologous gene product, on the one hand a lysosomal enzyme, on the other hand an IgG1-antibody, was expressed either under the control of the CMV-promoter or under the control of the hamster-ubiquitin/S27a-promoter. The two promoters were functionally linked to the CMV-enhancer. BGH polyA was used as the termination signal for the heterologous gene. The expression vectors which contained the CMV-promoter were based either on a modified pcDNA3 ("CMV¹", Invitrogen) or pBluescript vector ("CMV²";Stratagene) and additionally coded for the amplifiable selectable marker dihydrofolate-reductase. The expression vector with the hamster promoter on the other hand was based on the pAD-CMV vector (Werner et al., 1998). In order to express the heavy and light chain of the antibody, cotransfection was carried out with a second vector which contained a neomycin resistance gene as selectable marker. The CMV enhancer may, however, also be replaced by the SV40 enhancer.

By limited dilution in 96 well plates, after transfection, cell clones were selected and isolated in HT-free medium (with the addition of 400 µg/mL G418 in the case of co-transfection). Cell clones with the highest productivity with respect to the recombinant protein were subjected to stepwise DHFR-based gene amplification by increasing the methotrexate concentration step by step from 5 nM via 50 nM, 500 nM to 2 µM, in conjunction with a diluting cloning in each case. At each amplification stage about 20 to 30 clones with the highest productivity were selected.

Generally, the hamster promoter was found to have the highest performance. Both in the expression of the lysosomal enzyme and also in the expression of the antibody, productivities or titres were obtained which were 2 to 5 times

higher than those obtained with cells in which the heterologous gene was expressed under the control of the CMV promoter. Figure 1 shows by way of example the relative titres and relative specific productivities of the best cell clones at that particular amplification stage, the expression for the particular heterologous gene based on the CMV promoter being set at 1 (CMV¹ for the lysosomal enzyme, CMV² for the antibody).

Example 2: Isolation of high-expressing sICAM cells by GFP-based FACS sorting

The soluble form of the intercellular adhesion molecule ICAM1, sICAM, is a possible treatment for colds as it competes with the ICAM receptor for the binding of rhinoviruses and in this way can reduce or even prevent their interaction with the ICAM receptor, the prerequisite for entry into the cells and subsequent infection (Bella et al., 1999; Marlin et al., 1990).

sICAM was chosen as an example for the expression of a single-chained protein (480 amino acids) in CHO cells. For this CHO-DG44 were transfected with pBIDG-sICAM (Fig.3). The additional expression of GFP in pBIDG-sICAM transfected cells made it possible to use a FACS-based selection strategy. The therapeutic protein sICAM and GFP were jointly expressed by a bicistronic transcription unit and the DHFR by a separate transcription unit. Two to three weeks after the first selection in HT-free CHO-S-SFMII medium, 5% of the cells with the highest GFP fluorescence were sorted out. After about two weeks' cultivation the 5% cells with the highest GFP fluorescence were again isolated. This sequential sorting was carried out six times in all. A good correlation could be demonstrated between sICAM productivity and GFP fluorescence (Fig. 4). By FACS-assisted selection alone, without any MTX amplification step, cell pools with high specific productivities of up to 16 pg/cell/day were thus isolated in a very short time (Fig. 5). By combining the GFP-based selection with a single subsequent MTX amplification step it was even possible to increase productivity to above 30 pg/cell/day (Fig. 6). These productivities were achieved both with an amplification of a pool after the fourth sort with 500 nM MTX and also after

amplification of a pool after the sixth sort with 2 μ M MTX. In contrast to a stepwise amplification which usually starts with very low MTX concentrations in the range from 5 – 20 nM MTX, a higher MTX concentration had to be used from the outset to achieve an amplification effect. Thus, there was no significant increase in productivity as a result of the addition of 5 or 50 nM of MTX to cells from the fourth sorting or as a result of the addition of 500 nM of MTX to cells from the sixth sorting (Fig. 6). Obviously, the level of DHFT in the starting pools was already so high that total DHFR inhibition could only be achieved with a high dose of MTX. Moreover, the pre-sorted cell pools survived the selection phase much better in spite of the high initial dose of MTX, i.e. a cell population of high vitality was obtained in a shorter time than with the conventional step-by-step gene amplification strategy (Fig. 7).

Example 3: Isolation of cells with high expression of the mAb F19 by GFP-based FACS sorting

In a co-transfection CHO-DG44 cells were transfected with the plasmid combination pBIDG-F19HC and pBIN-F19LC (Fig. 3). The expressed humanised antibody F19 is directed against the surface molecule FAP which is synthesised by reactiven stroma fibroblasts (cf also EP 0 953 639). In the vector configurations used the two protein chains of the antibody are expressed by their own vector, which additionally also codes for a DHFR or neomycin-phosphotransferase selectable marker in a separate transcription unit. In addition, another selectable marker, GFP, is contained in the vector pBIDG-F19HC. By the transcriptional linking of the expression of the GFP and the heavy chain by means of an IRES element, in the co-transfection of CHO-DG44 with the vectors pBIDG-F19HC/pBIN-F19LC, cells with a high expression of the antibody F19 could rapidly be isolated solely by selecting the cells with a high GFP content using sequential FACS sorting. For this, after a first two- to three-week selection of the transfected cell pools in HT-free CHO-S-SFMII medium with the addition of 400 μ g/mL of G418, the 5% of cells with the highest GFP fluorescence were sorted out by FACS. This sorting was carried out up to six times in all, leaving a cultivation period of about 2 weeks between each sorting. Astonishingly there was found to be a

good correlation between F19 productivity and GFP fluorescence (Fig. 8), although both protein chains were expressed from their own vector and in the GFP-based FACS sorting it was only possible to select for the expression of the heavy chain, as a result of its transcriptional coupling with GFP. The productivities could be increased to 10 pg/cell/day (Fig. 9) and further increased to an average of 37 pg/cell/day by a single subsequent MTX amplification step, starting from the cell pool of the fifth sorting, by adding 1000 nM MTX to the selection medium. Comparable data could also be obtained by functionally linking the hamster promoter with the SV40 enhancer instead of the CMV enhancer. At the same time, the development time for selecting high-producing cells in comparison with a conventional stepwise gene amplification strategy, which generally comprises four amplification stages, with increasing amounts of MTX, could be reduced by half to about 120 days, with a concomitant significant reduction in the development capacities and costs.

LITERATURE

Adam, M.A. et al., *J Virol* 1991, 65, 4985 - 4990

Ausubel, F.M. et al., *Current Protocols in molecular biology*. New York : Greene Publishing Associates and Wiley-Interscience. 1994 (updated)

Bella, J. et al., *J. Struct. Biol.* 1999, 128, 69 - 74

Bennett, R.P. et al., *BioTechniques* 1998, 24, 478- 482

Chalfie, M. et al., *Science* 1994, 263, 802 - 805

Chamov, S.M. et al., *Antibody Fusion Proteins*, Wiley-Liss Inc., 1999

Davies, M.V. et al., *J Virol* 1992, 66, 1924 –1932

Faisst, S. et al., *Nucleic Acids Research* 1992, 20, 3 –26

Gossen, M. et al., *Curr Opin Biotech* 1994, 5, 516 - 520

Gubin, A.N. et al., *Biochem Biophys Res Commun* 1997, 236, 347 –350

Haber, D.A. et al., *Somatic Cell Genetics* 1982, 8, 499 - 508

Harris et al., *Protein Purification : A Practical Approach*, Pickwood and Hames, eds., IRL Press, 1995

Hemann, C. et al., *DNA Cell Biol* 1994, 13 (4), 437 - 445

Huston, C. et al., *Proc Natl Acad Sci USA* 1988, 85 (16), 5879 - 5883

Hu, S. et al., *Cancer Res.* 1996, 56 (13), 3055 - 3061

Jang, S.K. et al., *J Virol* 1989, 63, 1651 –1660

Kaufman, R.J. et al., *J Mol Biol* 1982, 159, 601 - 621

Kaufman, R.J., *Methods in Enzymology* 1990, 185, 537 - 566

Kortt, A.A. et al., *Protein Engineering* 1997, 10 (4), 423 - 433

Levenson, V.V. et al., *Human Gene Therapy* 1998, 9, 1233 –1236

Lottspeich F. and Zorbas H. eds., Bioanalytic, Spektrum Akad. Verl., 1998

Lovejoy, B. et al., *Science* 1993, 259, 1288 - 1293

Marlin, S.D. et al., *Nature* 1990, 344, 70 - 77

Meng, Y.G. et al., *Gene* 2000, 242, 201 - 207

Monaco, L. et al., *Gene* 1996, 180, 145 - 150

Morgan, R.A. et al., *Nucleic Acids Research* 1992, 20, 1293 - 1299

Mosser, D.D. et al., *BioTechniques* 1997, 22, 150- 161

Nolan, G.P. et al., *Proc Natl Acad Sci USA* 1988, 85, 2603 –2607

Ohshima, Y. et al., *J Mol Biol* 1987, 195, 247 - 259

Pack, P. et al., *Biotechnology* 1993, 11, 1271 – 1277

Pack, P. et al., *J Mol Biol* 1995, 246 (11), 28 - 34

Pelletier, J. et al., *Nature* 1988, 334, 320 –325

Perisic, O. et al., *Structure* 1994, 2, 1217 - 1226

Primig, M. et al., *Gene* 1998, 215, 181 –189

Ramesh, N. et al., *Nucleic Acids Research* 1996, 24, 2697 –2700

Sambrook, J., Fritsch, E.F. & Maniatis, T., Molecular Cloning: A Laboratory Manual
Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 1989

Scopes, R., Protein Purification, Springer Verlag, 1988

Simonson, C.C. et al., *Proc Natl Acad Sci USA* 1983, 80, 2495 - 2499

Sugimoto et al., *Biotechnology* 1994, 12, 694 - 698

Urlaub, G. et al., *Cell* 1983, 33, 405 - 412

Werner, R.G. et al., *Arzneim.-Forsch./Drug.Res.* 1998, 48, 870 - 880

Wigler, M. et al., *Proc Natl Acad Sci USA* 1980, 77, 3567 – 3570

SEQUENCE LISTING

<110> BOEHRINGER INGELHEIM PHARMA KG

<120> EXPRESSION VECTOR, METHODS FOR THE PRODUCTION OF HETEROLOGOUS GENE PRODUCTS AND FOR THE SELECTION OF RECOMBINANT CELLS PRODUCING HIGH LEVELS OF SUCH PRODUCTS

<130> Case 1-1412

<140>

<141>

<160> 1

<170> PatentIn Ver. 2.1

<210> 1

<211> 2406

<212> DNA

<213> Cricetulus griseus

<300>

<310> PCT/EP/96/04631

<311> 1996-10-24

<312> 1997-05-01

<400> 1

gatctccagg acagccatgg ctattacaca gagaaaccct gtctggaaaa acaaaaaatt
60

agtgtccatg tgtaaatgtg tggagtatgc ttgtcatgcc acatacagag gtagagggca
120

gtttatggga gtcagttcct attcttcctt tatggggac ctggggactg aactcaggc
180

atcaggcttg gcagaaagtg cattagctca cggagcctta tcattggcga aagctctc
240

aagtagaaaa tcaatgtgtt tgctcatagt gcaatcatta tgtttcgaga ggggaagggt
300

acaatcgttg gggcatgtgt ggtcacatct gaatagcagt agctccctag gagaattcca
360

agttcttgg tggtgtatca atgccttaa aggggtcaac aactttttt ccctctgaca
420

aaactatctt cttatgtcct tgtccctcat atttgaagta ttttattctt tgcagtgttg
480

aatatcaatt ctagcacctc agacatgtta ggtaagtacc ctacaactca ggttaactaa
540

tttaatttaa ctaatttaac cccaacactt tttctttgtt tatccacatt tgtggagtgt
600

gtgtgtgtgt gtgtgtgtgt gtgtgtgtgt gtgtgtgtgt gtgtgtgtgc
660

gcgcgcgcgc gcgcgtcgat cattctacct tttgtttaaa aaatgttagt ccaggggtgg
720

ggtgcactgt gaaagtctga gggtaacttg ctggggtcag ttctttccac tataggacag
780

aactccaggt gtcaactctt tactgacaga accatccaaa tagccctatc taattttagt
840

tttttattta tttatTTTTT gtttttcgag acagggtttc tctgtggcTT tggaggctgt
900

cctggaacta gctctttag accaggctgg ttcgaactc agagatccac ctgcctctgc
960
ctcctgagtg ctgggattaa aggcatgcmc caccacgct tggctctacc taatttaaa
1020
agagattgtg tgcacaagg gtgtcatgtc gccctgcaac caccgggggg caaaaaaaa
1080
aaaaaaaaaaa acttcactga agctgaagca cgatgatttgg tttactctgg ctggccaatg
1140
agctctagg agtctcctgt caaacagaat ctcaacaggc gcagcagtct tttttaaagt
1200
ggggttacaa cacaggttt tgcatatcatg gcattttatc taagctatcc cccagccaaa
1260
aatgtgtatt ttggaggcag cagagctaat agataaaaat gagggaagag cccacacagg
1320
ttatttagaa gataagcatc ttctttatataaaac caaaccacaaac tggaggaggt
1380
ctacccttag ggatggaaga aaagacattt agagggtgca atagaaaggg cactgagtt
1440
gtgaggtgga ggactgggag agggcgcaac cgcttaact gtcctgttt gcctatcc
1500
tggggacagc acatgttcctt atttttccca ggatgggcaa tctccacgtc caaacttgcg
1560
gtcgaggact acagtcattt tgcaaggttt cttactgtat ggctttaaa acgtgcaaaag
1620
gtgaccatta accgtttcac gctgggaggg cacgtgcggc tcagatgctt cctctgactg
1680
agggccagga gggggctaca cggaaagaggc cacacccgca cttgggaaga ctcgatttgg
1740
gcttcagctg gctgagacgc cccagcaggc tcctcggcta caccttcagc cccgaatgcc
1800
ttccggccca taacccttcc cttctaggca tttccggcga ggacccaccc tcgcgcacaaa
1860
cattcggccc catccccccgg tcctcacctg aatctctaac tctgactcca gagtttagag
1920
actataacca gatagccccgg atgtgtggaa ctgcacatcttgg ggacgagtag ttttagcaaa
1980
aagaaagcga cgaaaaacta caattcccgac agacttgtt gttacctctc ttctcatgct
2040
aaacaagccc cctttaaagg aaagccccctc ttagtcgcat cgactgtgtaa agaaaggcgt
2100
ttgaaacatt ttaatgttgg gcacaccgtt tcgaggaccg aaatgagaaaa gagcataggg
2160
aaacggagcg cccgagctag tctggactg cgtagacag ccgcggcgt tgcaacgggc
2220
aggcacttgc gtggacgcct aaggggcggg tcttcggcc gggaaagcccc gttggccgc
2280
gcggctcttc cttccgatc cggccatccgt ggtgagtg tgctgcgggc tgccgctccg
2340
gcttggggct tcccgctcg ctctcaccct ggtcggcggc tctaattccgt ctctttcga
2400
atgttag
2406

PATENT CLAIMS

1. Expression vector comprising a gene which codes for a protein of interest, functionally linked to a hamster-ubiquitin/S27a-promoter and a gene which codes for a fluorescent protein.
2. Expression vector according to claim 1, characterised in that it contains an amplifiable selectable marker gene.
3. Expression vector according to claim 1 or 2, characterised in that it comprises one or more enhancers functionally linked to the promoter or promoters.
4. Expression vector according to one of Claims 1 to 3, characterised in that it additionally contains an internal ribosomal entry site (IRES) which allows bicistronic expression of the gene which codes for a fluorescent protein, and of the gene which codes for a protein/product of interest.
5. Expression vector according to one of Claims 1 to 4, characterised in that the gene which codes for a fluorescent protein and the amplifiable selectable marker gene are located in one or in two separate transcription units.
6. Expression vector according to one of Claims 1 to 5, characterised in that the functional linking does not take place via intron sequences.
7. Expression vector according to one of Claims 1 to 6, characterised in that the amplifiable selectable marker gene codes for dihydrofolate-reductase (DHFR) or a fusion protein of the fluorescent protein and DHFR.
8. Expression vector according to one of Claims 3 to 7, characterised in that the enhancer is a CMV or SV40 enhancer.

9. Expression vector according to one of Claims 1 to 8, characterised in that it additionally comprises at least one polyadenylation signal.
10. Expression vector according to one of Claims 1 to 9, characterised in that instead of the gene which codes for a protein/product of interest, it comprises a multiple cloning site for the incorporation of a gene which codes for a protein of interest.
11. Eukaryotic host cell which has been transfected with an expression vector according to one of Claims 1 to 10.
12. Host cell according to claim 11, characterised in that it is a mammalian cell, preferably a CHO cell.
13. Host cell according to claim 11 or 12, characterised in that it has additionally been transfected with one or more vectors with genes which codes for one or more other proteins /products of interest and at least one other selectable marker.
14. Process for preparing a heterologous gene product, wherein a host cell according to one of Claims 11 to 13 is cultivated under conditions which allow expression of the gene product, and the gene product is isolated from the culture or culture medium.
15. Process for preparing a heteromeric protein/product, wherein a host cell according to claim 13, which has been co-transfected with expression vectors which code for different sub-units of the heteromeric protein/product, is cultivated under conditions which allow expression of the heteromeric protein/product, and the heteromeric protein/product is isolated from the culture or culture medium.
16. Process according to claim 15, characterised in that the heteromeric protein is an antibody.

17. Process according to one of claims 14 to 16, characterised in that the host cell is subjected to one or more gene amplification steps.
18. Process according to claim 17, characterised in that the amplifiable selectable marker is DHFR and the amplifying agent is methotrexate.
19. Process according to claim 18, characterised in that the host cell is subjected to only one gene amplification step with methotrexate.
20. Process according to one of claims 14 to 19, characterised in that it is carried out in a serum-free culture medium.
21. Process according to one of claims 14 to 20, characterised in that the host cell is cultivated in suspension culture.
22. Process for selecting a host cell which expresses a protein/product of interest, wherein a population of host cells according to one of claims 11 to 13 is cultivated under conditions which allow expression of the protein/product of interest and of the fluorescent protein, and the cell/cells which show the highest expression levels of fluorescent protein are identified and/or selected.
23. Process according to claim 22, characterised in that the selection is carried out using a Fluorescence-Activated Cell Sorter (FACS).
24. Process according to claim 22 or 23, characterised in that the correspondingly selected host cells are subjected to one or more additional gene amplification steps.
25. Process according to claim 24, characterised in that the amplifiable selectable marker is DHFR and the amplifying agent is methotrexate.

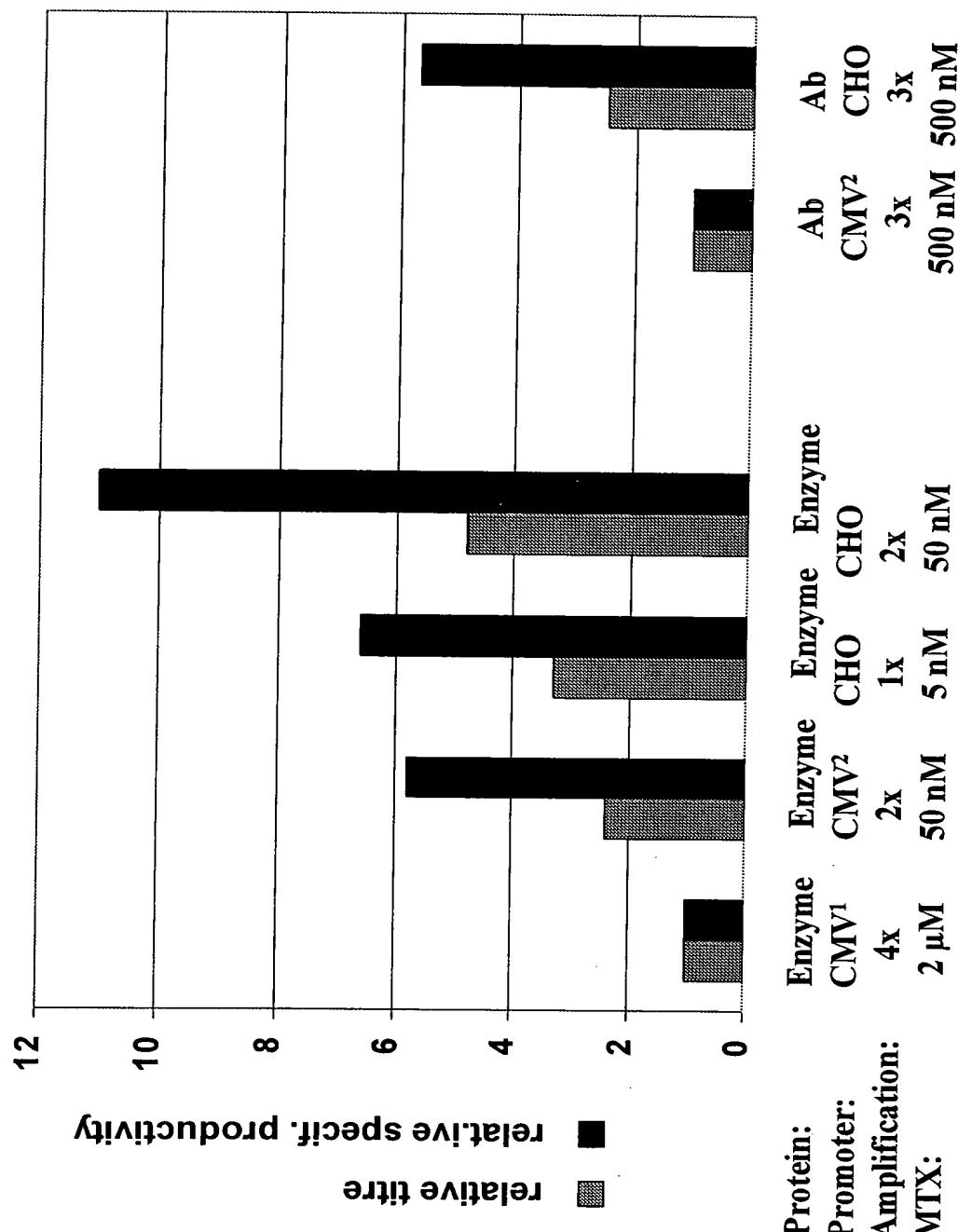


Figure 1

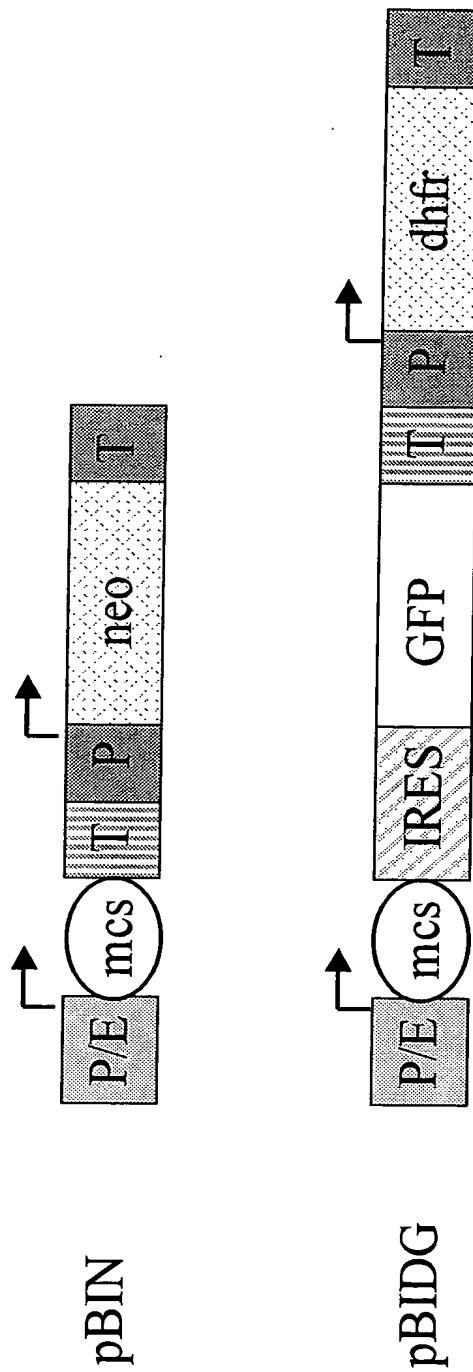
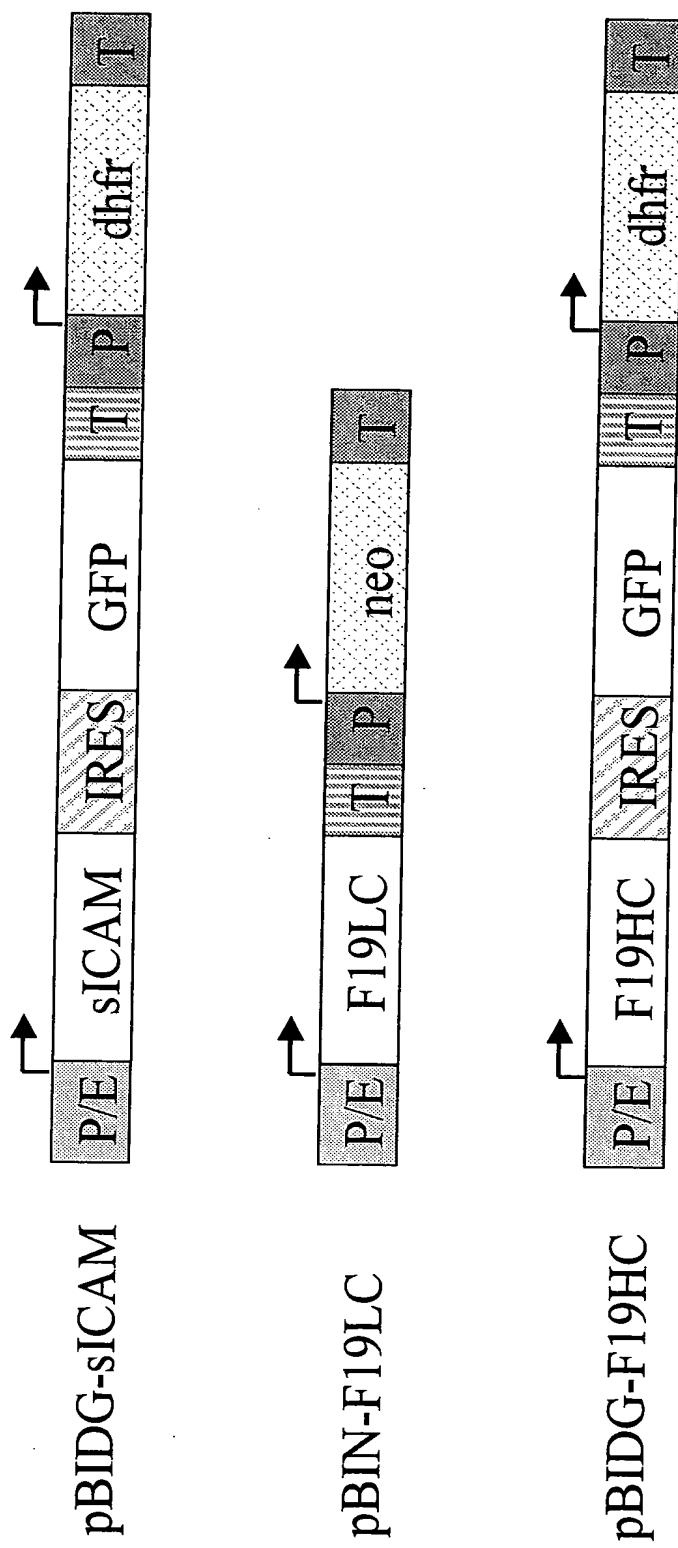


Figure 2

**Figure 3**

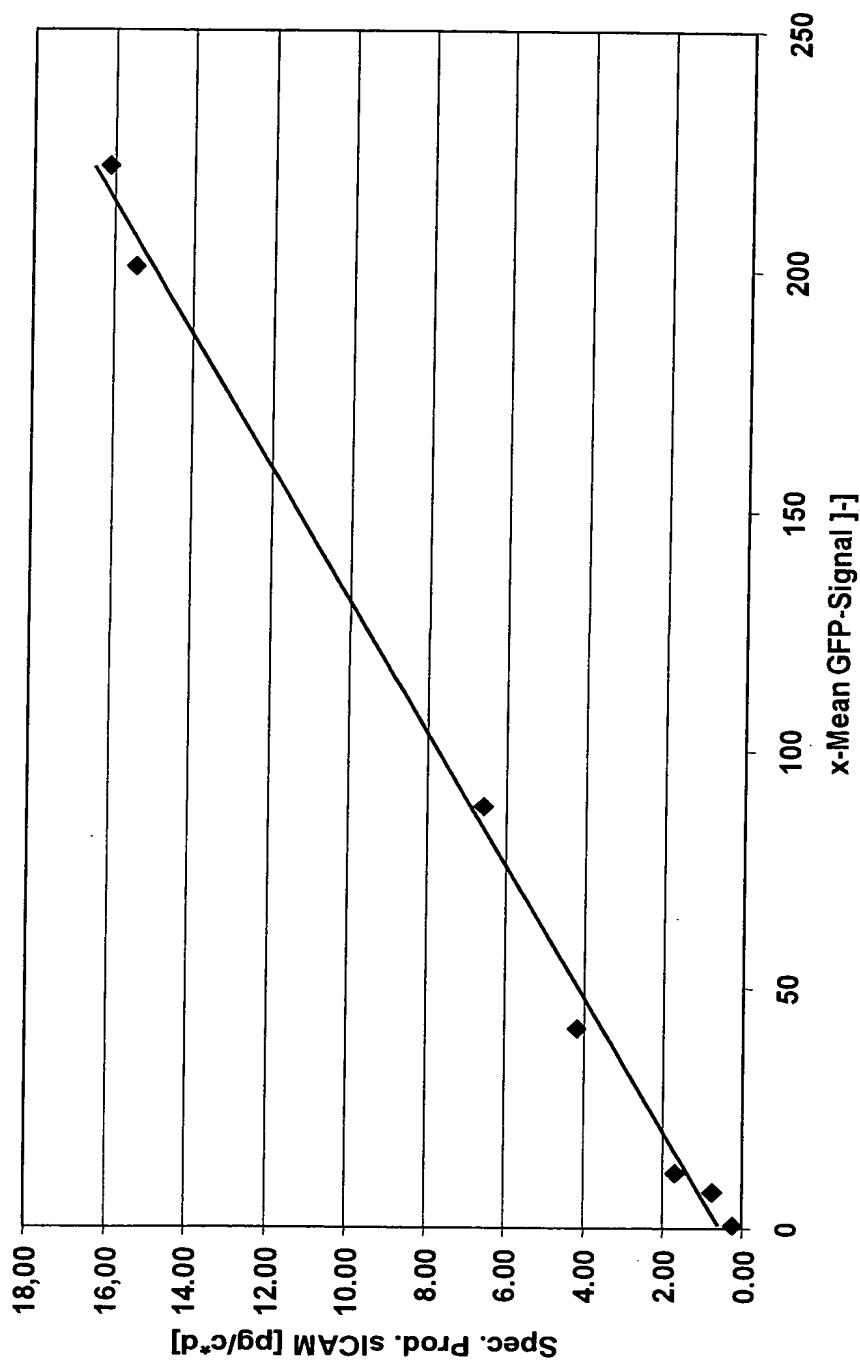


Figure 4

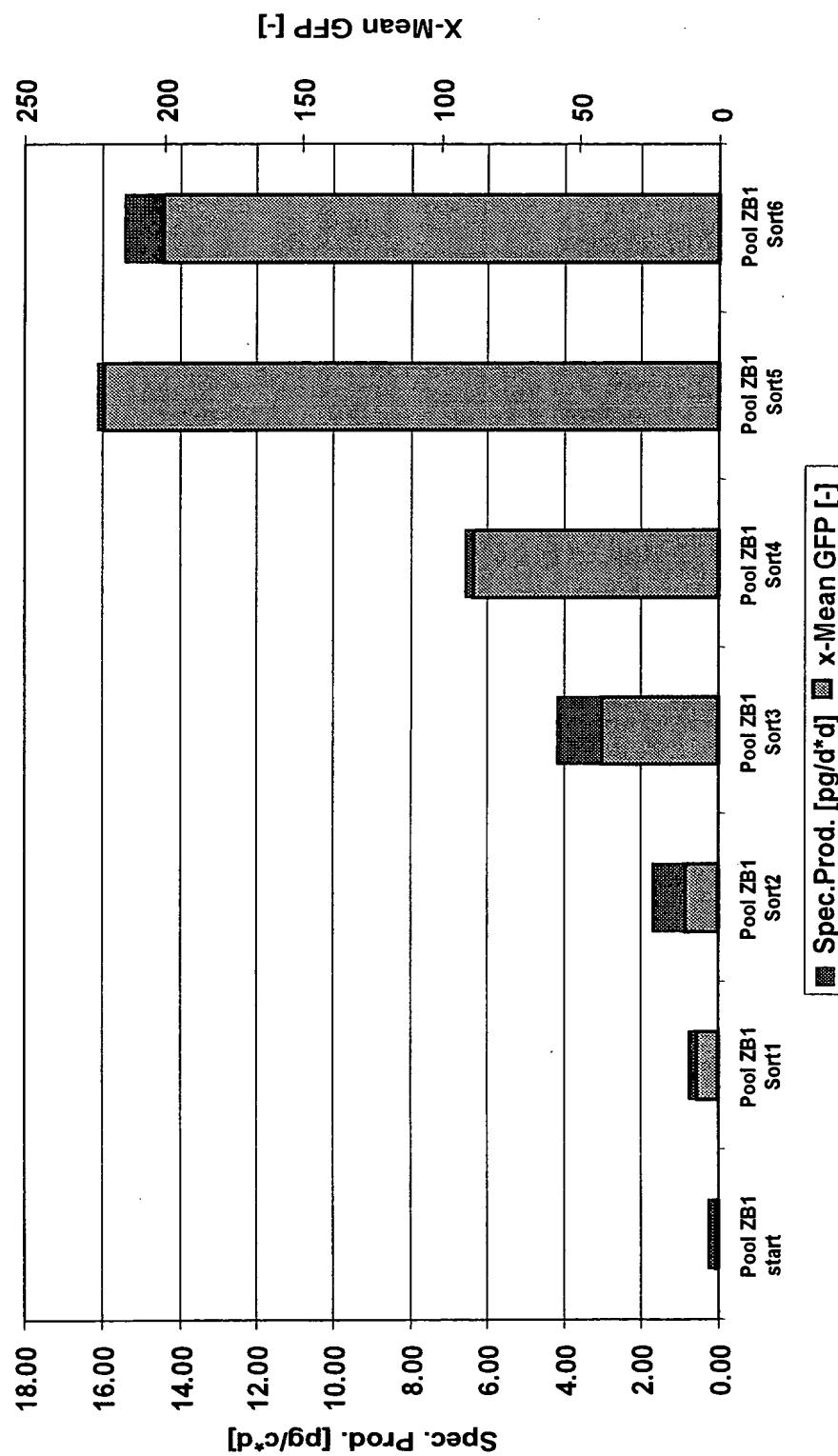


Figure 5

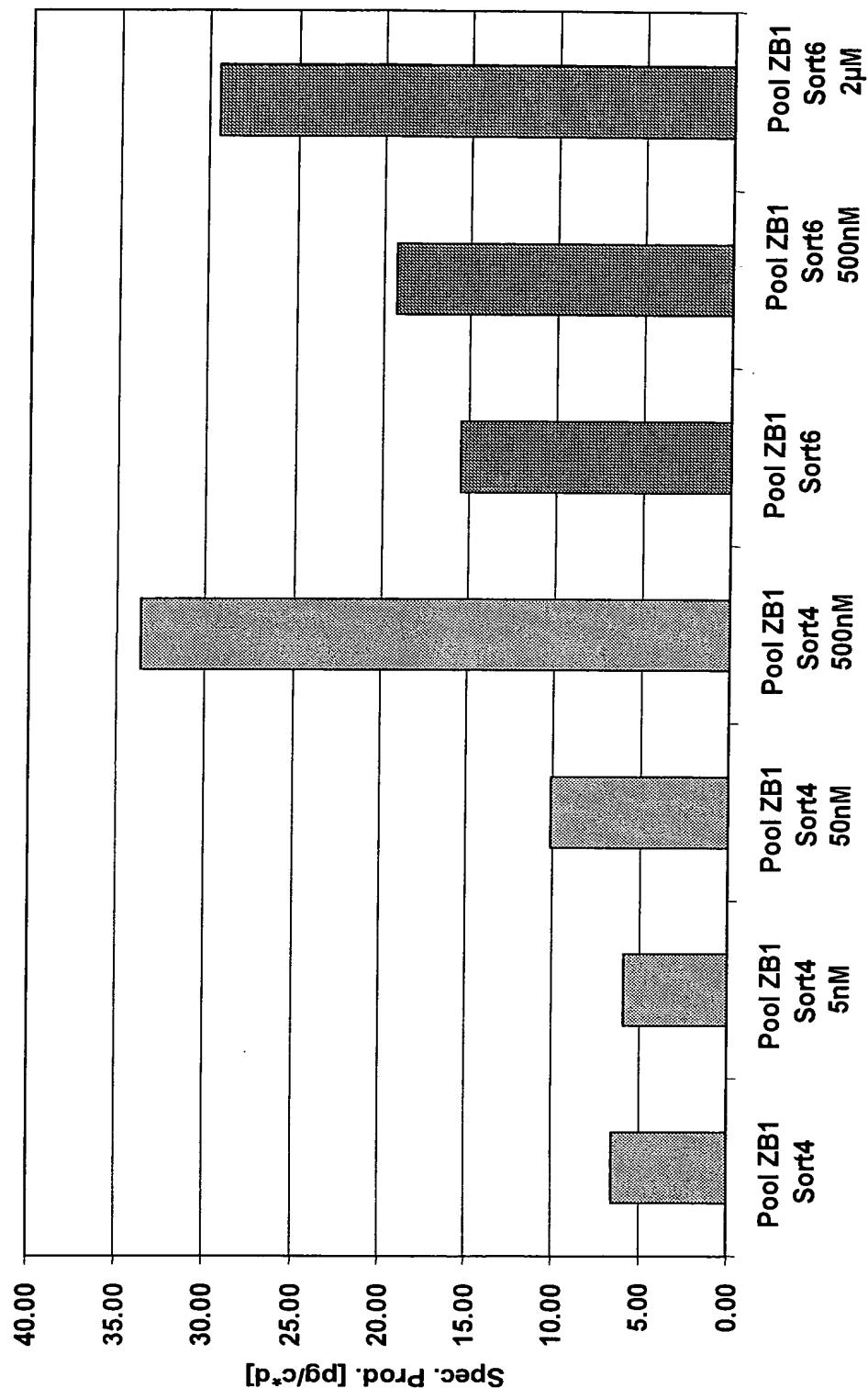


Figure 6

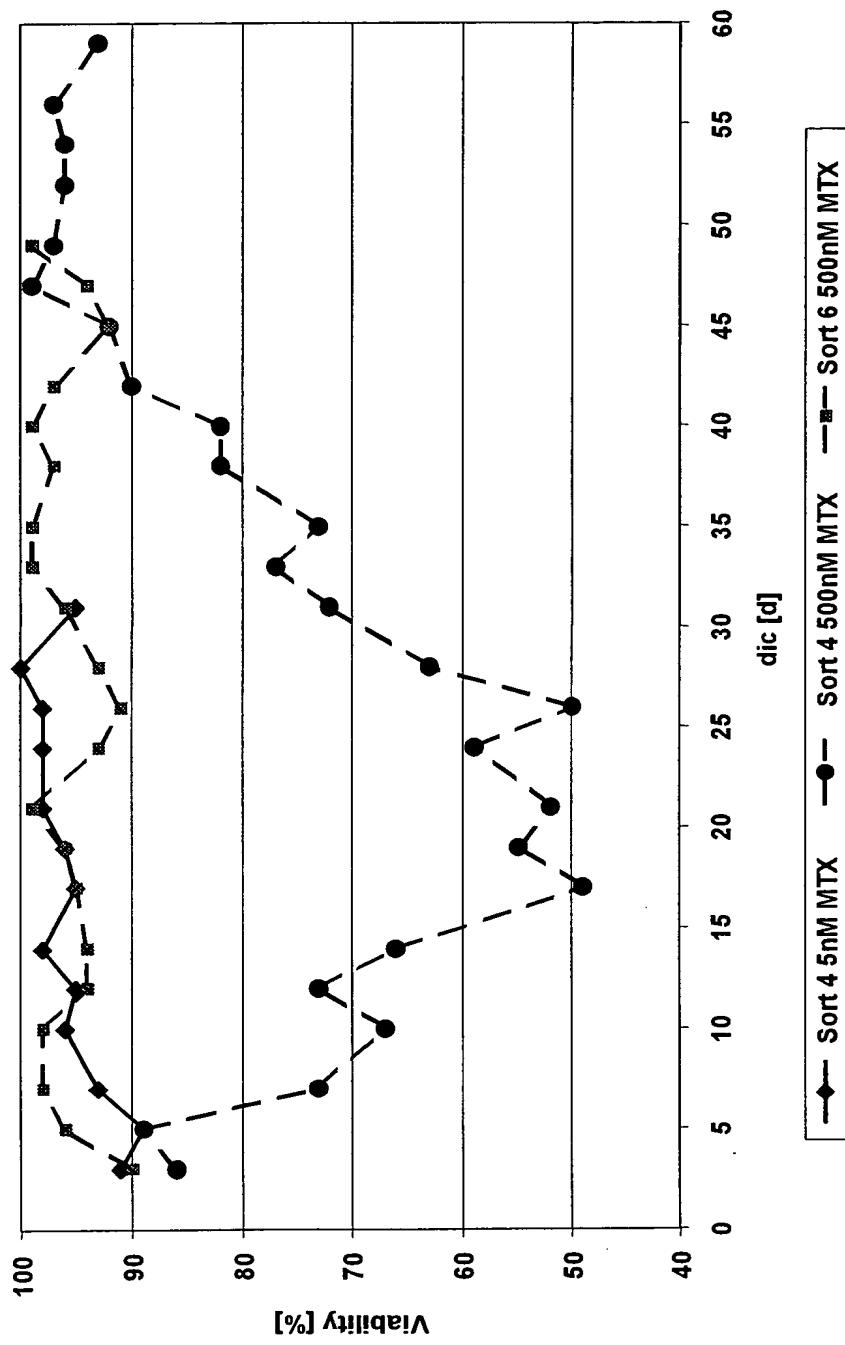


Figure 7

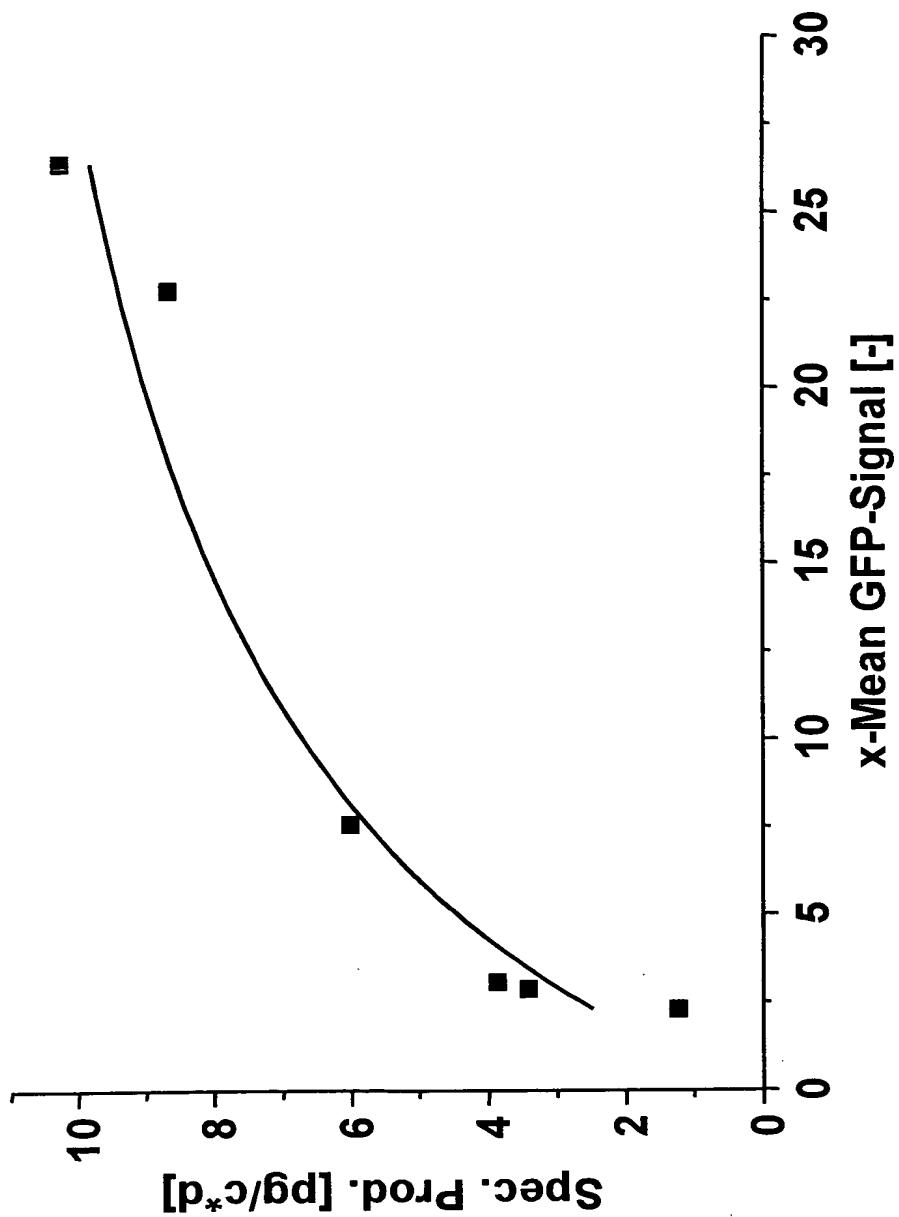


Figure 8

9/9

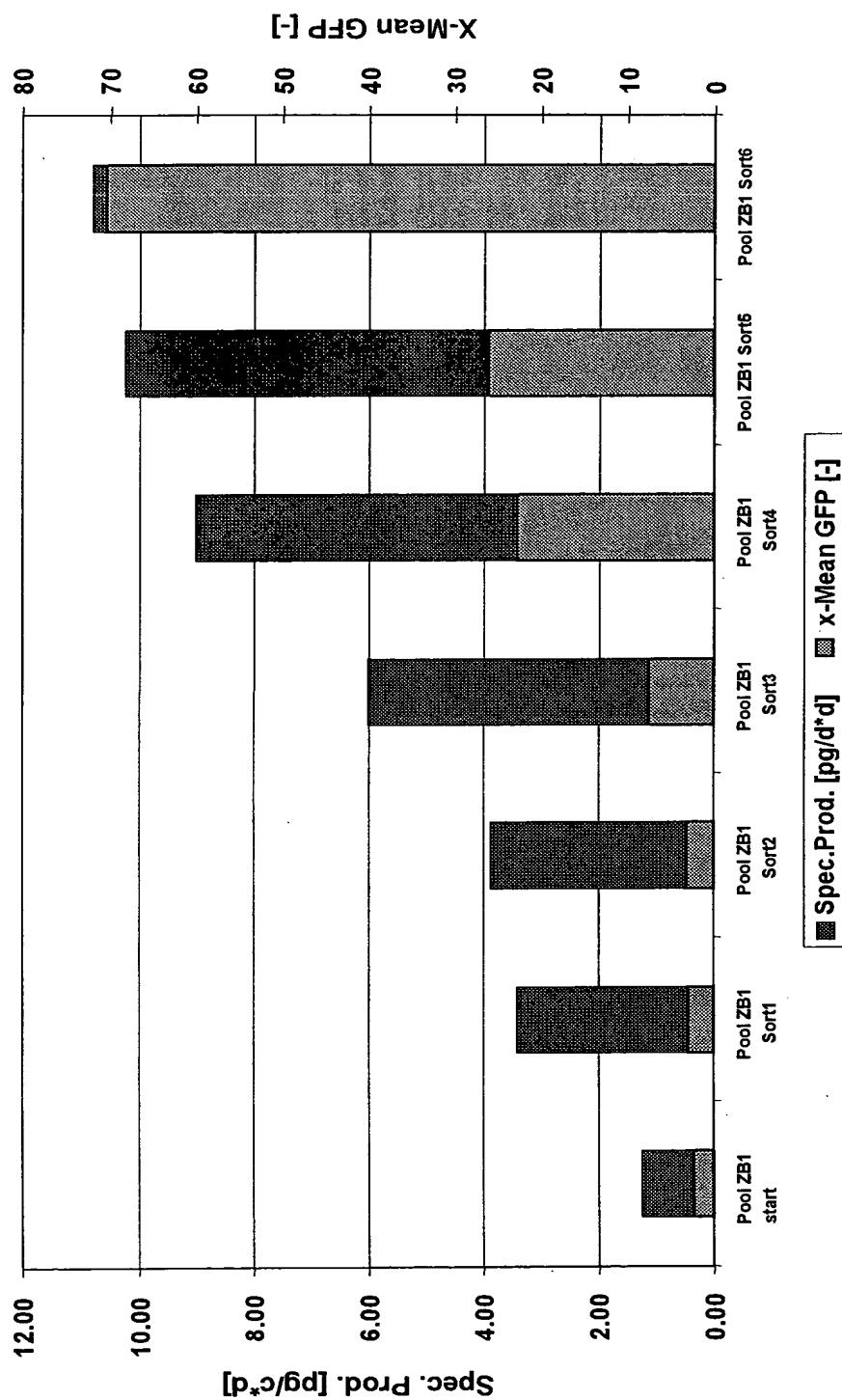


Figure 9

ABSTRACT

EXPRESSION VECTOR, METHODS FOR THE PRODUCTION OF
HETEROLOGOUS GENE PRODUCTS AND FOR THE SELECTION OF
RECOMBINANT CELLS PRODUCING HIGH LEVELS OF SUCH
PRODUCTS

An expression vector for eukaryotic cells comprising a gene which codes for a protein of interest, functionally linked to a hamster-ubiquitin/S27a-promoter and a gene which codes for a fluorescent protein. Preferably the expression vector also contains an amplifiable selectable marker gene.

The invention also describes host cells, preferably mammalian cells, which have been transfected with the expression vector, processes for producing heterologous gene products and a method of selecting high-producing cells.

BUNDESREPUBLIK DEUTSCHLAND



Prioritätsbescheinigung über die Einreichung einer Patentanmeldung

Aktenzeichen: 102 56 083.8

Anmeldetag: 29. November 2002

Anmelder/Inhaber: Boehringer Ingelheim Pharma GmbH & Co KG,
Ingelheim/DE

(vormals: Boehringer Ingelheim Pharma KG)

Bezeichnung: Expressionsvektor, Verfahren zur Herstellung von
heterologen Genprodukten und Selektionsverfahren
für hochproduzierende rekombinante Zellen

IPC: C 12 N 15/85

Die angehefteten Stücke sind eine richtige und genaue Wiedergabe der ursprünglichen Unterlagen dieser Patentanmeldung.

München, den 24. September 2003

Deutsches Patent- und Markenamt

Der Präsident

Im Auftrag

A handwritten signature in black ink, appearing to read "Brosig".

EXPRESSIONSVEKTOR, VERFAHREN ZUR HERSTELLUNG VON
HETEROLOGEN GENPRODUKTEN UND SELEKTIONSSVERFAHREN FÜR
HOCHPRODUZIERENDE REKOMBINANTE ZELLEN

5 Gebiet der Erfindung

Die Erfindung betrifft ein Selektionsverfahren für hochproduzierende rekombinante Zellen, ein Verfahren zur Herstellung von heterologen Genprodukten sowie Expressionsvektoren und damit transfizierte Wirtszellen, die in diesen Verfahren 10 verwendbar sind.

 Hintergrund der Erfindung

Säugerzellen sind die bevorzugten Wirtszellen zur Produktion komplexer 15 biopharmazeutischer Proteine, da die post-translational durchgeführten Modifikationen sowohl in funktionaler als auch in pharmakokinetischer Hinsicht humankompatibel sind. Kommerziell relevante Zelltypen sind Hybridomas, Myelomas, CHO (Chinese Hamster Ovary)-Zellen und BHK (Baby Hamster Kidney)-Zellen. Die Kultivierung der Wirtszellen erfolgt zunehmend unter serum- und 20 proteinfreien Produktionsbedingungen. Gründe hierfür sind die damit verbundene Kostenreduktion, die geringere Interferenz bei der Aufreinigung des rekombinanten Proteins sowie die Reduktion des Potentials zur Einführung von Pathogenen (z.B. 25 Prionen, Viren). Der Einsatz von CHO-Zellen als Wirtszellen findet dabei immer mehr Verbreitung, da diese Zellen sich an Suspensionswachstum in serum- und proteinfreiem Medium adaptieren lassen und sie zudem von den regulatorischen Behörden als sichere Produktionszellen angesehen und akzeptiert werden.

Zur Erzeugung einer stabilen Säugerzelllinie, die ein heterologes Gen von Interesse 30 exprimiert, wird das heterologe Gen in der Regel gemeinsam mit einem selektierbaren Markergen, wie z.B. Neomycin-Phosphotransferase, durch Transfektion in die gewünschten Zelllinie eingebracht. Das heterologe Gen und das selektierbare Markergen können dabei entweder gemeinsam von einem einzelnen Vektor oder von jeweils separaten Vektoren, die co-transfiziert werden, exprimiert

werden. Zwei bis drei Tage nach der Transfektion werden die Zellen in Medium überführt, das ein Selektionsmittel enthält, z.B. G418 bei Verwendung des Neomycin-Phosphotransferase-Gens, und für einige Wochen unter diesen selektiven Bedingungen kultiviert. Die hochwachsenden resistenten Zellen können dann isoliert

5 und hinsichtlich der Expression des gewünschten Genprodukts untersucht werden. Bedingt durch die willkürliche und ungerichtete Integration in das Wirtszellgenom erhält man eine Population von Zellen, die völlig unterschiedliche Expressionsraten des heterologen Gens aufweisen. Darunter können auch nicht-exprimierende Zellen sein, bei denen zwar der Selektionsmarker exprimiert wird, aber nicht das Gen von

10 Interesse. Zur Identifizierung von Zellklonen, die eine sehr hohe Expression des heterologen Gens von Interesse aufweisen, muss deshalb eine Vielzahl von Klonen überprüft und getestet werden, resultierend in einem hohen Zeit-, Arbeits- und Kostenaufwand.

15 Genamplifikation ist ein weitverbreitetes Phänomen in tierischen Zellkulturen, das für die Produktion rekombinanter biopharmazeutischer Proteine genutzt wird. Durch die Genamplifikation wird die ursprünglich relativ geringe Produktivität vieler Säugerzelllinien drastisch verbessert. Eine weit verbreitete Amplifikationstechnik stellt das Dihydrofolat-Reduktase (DHFR)-basierte Genamplifikationssystem dar, das

20 sehr häufig in DHFR-defizienten Chinese Hamster Ovary (CHO)-Zellen Verwendung findet. Dabei werden die DHFR-defizienten CHO-Zellen, z.B. CHO-DUKX (ATCC CRL-9096) oder CHO-DG44 (Urlaub et al., 1983), mit einem geeigneten Vektorsystem, das für DHFR und das Protein von Interesse kodiert, transfiziert. Daraufhin werden die Transfektanten in einem Medium ohne Glycin, Hypoxanthin

25 und Thymidin selektiert. Die Amplifikation und somit die Etablierung von hochproduzierenden Zelllinien wird über die steigende Zugabe von Methotrexat (MTX), einem Inhibitor der Dihydrofolat-Reduktase, erreicht (Kaufman et al., 1982; US 4,656,134). Die anschließende Selektion der erhaltenen hochproduzierenden Zellen unterliegt auch hierbei dem Zufallsprinzip und basiert auf

30 Wahrscheinlichkeiten, wodurch dieser Selektionsschritt überaus arbeits- und zeitintensiv ist.

Zur besseren und schnelleren Verfolgung der Gentransformation und -expression wurden verschiedenste Methoden entwickelt. Diese beinhalteten zunächst die Verwendung von Reportermolekülen wie z.B. Chloramphenicol-Acetyltransferase, Luciferase, β -Galactosidase oder von Fusionsproteinen, die die Kodierregionen von β -Galactosidase oder Luciferase enthalten. Der Nachteil dieser entsprechenden Reportergen-Assays besteht allerdings darin, dass die Zellen fixiert oder lysiert und mit exogen hinzugefügten Substraten und Co-Faktoren inkubiert werden müssen. Eine Weiterkultivierung der analysierten Zellen ist somit ausgeschlossen. Eine neuere Methode, basierend auf der Co-Expression des *E.coli* Enzyms β -Galactosidase, ermöglicht zwar eine Sortierung von lebenden Zellen mittels eines FACS-Geräts (Nolan et al., 1988), allerdings ist hier zur Beladung der Zellen mit dem fluorogenen Substrat eine hypotonische Vorbehandlung erforderlich. Diese Aktivität muss vor dem FACS-basierten Sortierung noch inhibiert werden.

Mit der Einführung des grünen fluoreszierenden Proteins (GFP) aus *Aequorea victoria* und den daraus entwickelten GFP-Mutanten als Reportermolekül wurde die Identifizierung von Zellen, die ein heterologes Gen exprimieren, erheblich erleichtert. Die Co-Expression von GFP ermöglichte eine *in vivo* Echtzeit-Analyse und das Sortieren von Transfektanten auf Basis ihrer Fluoreszenz, ohne dass zusätzliche Substrate oder Co-Faktoren benötigt werden. Der Einsatz von GFP als Reportermolekül zur Verfolgung des Gentransfers wurde in verschiedenen Publikationen beschrieben. Chalfie et al. beschreiben in den Patenten US 5,491,084 und US 6,146,826 eine Methode zur Selektion von Zellen, die ein Protein von Interesse exprimieren. Diese Methode beinhaltet die Co-Transfektion von Zellen durch ein DNA-Molekül, das die kodierende Sequenz für das Protein von Interesse enthält, und einem zweiten DNA-Molekül, das das GFP-Gen kodiert. Nachfolgend werden die GFP-exprimierenden Zellen selektiert. Gubin et al. (1997) untersuchten die Stabilität der GFP-Expression in CHO-Zellen in Abwesenheit von selektiven Wachstumsbedingungen. Die Zellen wurden dabei mit einem Plasmid, das sowohl GFP als auch Neomycin-Phosphotransferase enthielt, transfiziert. Mosser et al. (1997) verwendeten zur Identifizierung und Selektion von Zellen, die induzierbares Produkt exprimierten, ein Plasmid, das eine bicistronische Expressionskassette,

kodierend für GFP und ein Zielgen (oder auch Gen von Interesse genannt), enthielt. Das Zielgen stand dabei unter der Kontrolle eines regulierbaren Promotors. Die Kopplung der GFP- und der Zielgenexpression wurde durch die Verwendung eines viralen IRES („Internal Ribosome Entry Site“)-Elements erreicht, wodurch eine 5 bicistronische mRNA, die für GFP und das Protein von Interesse kodierte, exprimiert wurde. Das hierbei verwendete Plasmid enthielt selbst kein selektierbares Markergen. Dieses wurde deshalb durch ein zweites Plasmid in einer Co-Transfektion oder in einer nachfolgenden Transfektion eingebracht. Hingegen verwendeten Levenson et al. (1998) retrovirale Vektoren mit einer bicistronischen 10 Expressionskassette, in der vor der IRES-Sequenz das Gen von Interesse kloniert werden kann. Die auf die IRES-Sequenz folgende Sequenz kodierte hingegen für ein selektionierbares Markergen, wobei es sich dabei entweder um einen Marker handelte, der Resistenz gegen G418, Puromycin, Hygromycin B, Histidinol D oder Phleomycin vermittelte, oder um GFP.

15

Es wurden auch bereits Vektoren beschrieben, die ein IRES-Element aus der Familie der Picornaviren enthielten, wobei das IRES-Element zwischen dem Produktgen und einem selektionierbaren Markergen positioniert war (Pelletier et al., 1988; Jang et al., 1989; Davies et al., 1992).

20

GFP wurde auch erfolgreich mit Resistenzmarkergenen fusioniert. Zum Beispiel beschreibt Bennett et al. (1998) ein GFP/Zeomycin-Fusionsprotein. Dieser bifunktionale Selektionsmarker konnte erfolgreich zur Identifizierung und Selektion von transfizierten Säugerzellen eingesetzt werden. Primig et al. (1998) hingegen 25 setzten ein Fusionsprotein aus GFP und Neomycin-Phosphotransferase für ihre Enhancer-Studien ein.

In der Publikation von Meng et al. (2000) und in der internationalen Patentanmeldung WO 01/04306 wurde zur Selektion und Identifizierung von Zellen mit hoher 30 Expression eines rekombinanten Proteins ein Expressionssystem eingesetzt, in dem das Gen von Interesse zusammen mit dem amplifizierbaren Selektionsmarkergen DHFR und einem GFP-Gen von einem einzigen Vektor aus exprimiert wurde. Die drei Gene waren dabei entweder in einer Transkriptionseinheit zusammengefasst

oder auf zwei Einheiten verteilt. Durch diese räumliche und transkriptionelle Verknüpfung aller drei Gene in einem einzigen Expressionsvektor sollte deren Wahrscheinlichkeit einer Co-Amplifikation unter Selektionsdruck erhöht werden und somit Hochproduzenten-Klone identifiziert und selektiert werden. Die besten 5 Klone, die durch Anwendung der kombinierten Selektion mittels amplifizierbarem DHFR-Selektionsmarker und GFP-basierter FACS-Sortierung isoliert wurden, exprimierten das Protein von Interesse in einer Größenordnung von maximal 3 bis 10 4.5 pg pro Zelle und Tag. Dabei wurden die Experimente mit adhärenten Zellen und in serumhaltigem Medium durchgeführt, also mit Zellen und unter Bedingungen, die bekanntermaßen wesentlich robuster und durch höhere Grundproduktivitäten gekennzeichnet sind.

Zusammenfassung der Erfindung

15 Aufgabe der vorliegenden Erfindung war daher die Entwicklung eines Selektions-
systems für rekombinante Zellen mit erhöhter Produktivität, das folgende
Anforderungen erfüllt:
20 (1) Verkürzung der Zeit für die Entwicklung von hochproduzierenden Zellen zur
Herstellung biopharmazeutischer Proteine bei gleichzeitiger Senkung der
Entwicklungskosten;
25 (2) Hoher Durchsatz bei der Selektion von hochproduzierenden Zellen bei geringem
Kapazitätsaufwand;
30 (3) Einsatz von „fermentationsrobusten“ hochproduzierenden Zellen, die z.B. eine
geringere Wachstumsbeeinträchtigung bei erhöhten Methotrexat-Konzentrationen
zeigen;
(4) Transfektion, Selektion und Kultivierung der suspensionsadaptierten Zellen
vorzugsweise in serumfreiem Medium;
(5) Reduktion der erforderlichen Genamplifikationsschritte.

30 Aufgabe der Erfindung war ferner die Bereitstellung von Expressionsvektoren und
damit transfizierten Wirtszellen, die in diesem Klonselektionssystem verwendbar

sind, sowie eines Verfahrens zur Herstellung von heterologen Genprodukten unter Verwendung dieser Wirtszellen.

Diese Aufgaben werden gemäß einem Aspekt der vorliegenden Erfindung mit Hilfe

5 eines Expressionsvektors gelöst, der ein Gen, das für ein Protein von Interesse kodiert (im folgenden auch: „Gen von Interesse“), in funktioneller Verknüpfung mit einem Hamster-Ubiquitin/S27a-Promotor und einem Gen, das für ein fluoreszierendes Protein kodiert, umfasst.

10 Der Expressionsvektor enthält vorzugsweise noch ein amplifizierbares Selektionsmarkergen, z.B. das Gen für die Dihydropholat-Reduktase (DHFR). Ferner

15 enthält ein bevorzugter Expressionsvektor weitere regulatorische Elemente, beispielsweise einen Enhancer in funktioneller Verknüpfung mit dem Promotor. Ferner enthält der Expressionsvektor vorzugsweise zusätzlich eine interne Ribosomenbindungsstelle (IRES), welche die bicistronische Expression des Gens, das für ein fluoreszierendes Protein kodiert, und des Gens von Interesse ermöglicht.

Die Erfindung betrifft auch Basisvektoren, die anstelle des Gens von Interesse eine multiple Klonierungsstelle zum Einbau eines solchen Gens aufweisen, d.h. einen

20 Sequenzbereich mit multiplen Schnittstellen für Restriktionsendonukleasen.

Ein weiterer Aspekt der vorliegenden Erfindung sind Wirtszellen, die mit einem der genannten Expressionsvektoren transfiziert worden sind. Hierbei handelt es sich um eukaryontische Wirtszellen, vorzugsweise Säugerzellen, wobei Nagerzellen wie

25 Hamsterzellen und insbesondere CHO-Zellen oder BHK-Zellen besonders bevorzugt sind.

Ein anderer Aspekt der vorliegenden Erfindung ist ein Verfahren zur Herstellung eines heterologen Genprodukts, bei dem eine mit dem erfindungsgemäßigen

30 Expressionsvektor transfizierte Wirtszelle unter Bedingungen, die eine Expression des Genprodukts ermöglichen, kultiviert und das Genprodukt aus der Kultur oder dem Kulturmedium isoliert wird.

In einer besonderen Ausführungsform der Erfindung wird die Wirtszelle mit dem erfindungsgemäßen Expressionsvektor sowie zusätzlich mit einem oder mehreren Vektoren mit Genen, die für ein oder mehrere andere Proteine von Interesse kodieren, transfiziert, bevorzugt co-transfiziert.

5

In diesem Zusammenhang stellt die vorliegende Erfindung ein Verfahren zur Herstellung eines heterodimeren Proteins zur Verfügung, bei dem man eine derartige Wirtszelle, die mit Expressionsvektoren, welche für unterschiedliche Untereinheiten des heterodimeren Proteins kodieren, co-transfiziert worden ist, unter Bedingungen, 10 die eine Expression des heterodimeren Proteins ermöglichen, kultiviert und das heterodimere Protein aus der Kultur oder dem Kulturmedium isoliert. Ein spezieller Anwendungsfall für ein solches Verfahren ist die Herstellung von Antikörpern und 15 deren Untereinheiten.

Ein weiterer Aspekt der vorliegenden Erfindung ist ein Verfahren zur Selektion einer Wirtszelle, die ein Protein von Interesse exprimiert, bei dem eine Population von Wirtszellen, die mit einem erfindungsgemäßen Expressionsvektor transfiziert worden sind, unter Bedingungen kultiviert wird, die eine Expression des Proteins von Interesse und des fluoreszierenden Proteins ermöglichen, und die Zelle(n) 20 identifiziert und/oder selektiert werden, welche die höchsten Expressionsraten an fluoreszierendem Protein zeigen. Die Selektion erfolgt dabei vorzugsweise mit Hilfe eines Fluoreszenz-aktivierten Zellsortiergeräts (FACS).

Es wurde überraschenderweise gefunden, dass mit dem erfindungsgemäß bereitgestellten System in kürzester Zeit Zellpools isoliert werden können, die ohne Genamplifikationsschritt durchschnittliche spezifische Produktivitäten von über 15 pg (einzelkettiges Protein) bzw. 10 pg (humanisierter Antikörper) rekombinantes Protein 25 pro Zelle und Tag exprimieren. Durch einen einzigen DHFR-basierten Genamplifikationsschritt konnten die spezifischen Produktivitäten noch auf über 30 pg pro 30 Zelle und Tag gesteigert werden. Die bei den Zellpools erzielten Produktivitäten liegen damit um den Faktor 8 bis 10 höher als die maximalen Produktivitäten der besten bisher publizierten Zellklone.

Erstaunlicherweise besteht auch eine sehr gute Korrelation zwischen der Expression des Proteins von Interesse und des fluoreszierenden Proteins. Diese ist sogar bei einer Co-Transfektion gegeben, wenn - wie im Fall eines exprimierten Antikörpers beide Immunglobulinketten jeweils von einem eigenen Vektor exprimiert werden und

5 bei der FACS-Sortierung nur auf die Expression der einen Kette, bedingt durch deren transkriptionelle Kopplung mit dem fluoreszierenden Protein, selektioniert werden kann. Die hohen Expressionsraten des fluoreszierenden Proteins haben dabei keinerlei negativen Einfluss auf Zellwachstum und -vitalität. Zudem kann die Entwicklungszeit zur Selektion von hochproduzierenden Zellen im Vergleich zu einer

10 konventionellen stufenweisen Genamplifikationsstrategie zumindest um die Hälfte reduziert werden, einhergehend mit einer signifikanten Reduktion der Entwicklungskapazitäten und -kosten.

15 Beschreibung der Abbildungen

Abbildung 1 zeigt einen Vergleich der erzielten Expressionslevel von rekombinanten Zellklonen, in denen das heterologe Genprodukt entweder unter der Kontrolle des CMV-Promotors oder unter der des Hamster-Ubiquitin/S27a-Promotors exprimiert wird. Beide Promotoren sind dabei funktionell mit dem CMV-Enhancer verknüpft und die Terminationssequenz, BGH poly A, ist in allen Fällen identisch. Im Fall von CMV¹ handelt es sich um einen pcDNA3- (Invitrogen, Kalsruhe, DE), in CMV² um einen pBluescript- (Stratagene, La Jolla, CA, US) und bei CHO um einen pAD-CMV-basierten Expressionsvektor (Werner et al., 1998). Bei der Expression des lysosomalen Enzyms enthalten sämtliche Expressionsvektoren den amplifizierbaren Selektionsmarker Dihydrofolat-Reduktase (DHFR) und die Expression des heterologen Gens wurde durch nachfolgende Amplifikationsschritte mit Methotrexat (MTX) gesteigert. Zur Expression der beiden Ketten des Antikörpers (Ak) wurde eine Co-Transfektion mit einem zweiten Vektor, der als Selektionsmarker ein Neomycin-Resistenzgen enthält, durchgeführt. Die erzielten Titer bzw. spezifischen Produktivitäten sind in Relation zur CMV Promotor-basierten Expression, die als 1 gesetzt wurde (CMV¹ für Enzym, CMV² für Ak), angegeben.

Abbildung 2 zeigt eine schematische Darstellung der Basisvektoren, die zur Expression der rekombinanten Proteine in CHO-DG44 Zellen verwendet wurden. Bei „P/E“ handelt es sich um eine Kombination aus CMV-Enhancer und Hamster-Ubiquitin/S27a-Promotor, bei „P“ lediglich um ein Promotorelement und bei „T“ um 5 ein Terminationssignal für die Transkription, das zur Polyadenylierung der transkribierten mRNA benötigt wird. Die Position und Richtung der Transkriptionsinitiation innerhalb jeder Transkriptionseinheit wird durch einen Pfeil angezeigt. Zur Klonierung der heterologen Gene ist nach dem Promotorelement ein Sequenzbereich mit multiplen Schnittstellen für Restriktionsendonukleasen (*multiple cloning sites* - mcs) eingefügt. Der amplifizierbare Selektionsmarker Dihydrofolat-10 Reduktase ist mit „dhfr“ und der Selektionsmarker Neomycin-Phosphotransferase mit „neo“ abgekürzt. Das aus dem Encephalomyocarditis-Virus stammende IRES-Element dient als interne Ribosomenbindungsstelle innerhalb der bicistronischen Transkriptionseinheit und ermöglicht die Translation des nachfolgenden grünen 15 fluoreszierenden Proteins „GFP“.

Abbildung 3 zeigt eine schematische Darstellung der eukaryontischen Expressionsvektoren, die jeweils für ein biopharmazeutisches Protein kodieren und zur Transfektion von CHO-DG44 Zellen eingesetzt wurden. Bei „P/E“ handelt es sich 20 um eine Kombination aus CMV-Enhancer und Hamster Ubiquitin/S27a Promotor, bei „P“ lediglich um ein Promotorelement und bei „T“ um ein Terminationssignal für die Transkription, das zur Polyadenylierung der transkribierten mRNA benötigt wird. Die Position und Richtung der Transkriptionsinitiation innerhalb jeder Transkriptionseinheit wird durch einen Pfeil angezeigt. Der amplifizierbare 25 Selektionsmarker Dihydrofolat-Reduktase ist mit „dhfr“ und der Selektionsmarker Neomycin-Phosphotransferase mit „neo“ abgekürzt. Das aus dem Encephalomyocarditis-Virus stammende IRES-Element dient als interne Ribosomenbindungsstelle innerhalb der bicistronischen Transkriptionseinheit und ermöglicht die Translation des nachfolgenden grünen fluoreszierenden Proteins „GFP“. „sICAM“ kodiert für das lösliche intrazelluläre Adhäsionsmolekül (US 30 5,412,216), wohingegen „F19HC“ und „F19LC“ für die schwere bzw. leichte Kette des humanisierten Antikörpers F19 kodieren (EP 953 639).

Abbildung 4 zeigt die Korrelation zwischen der sICAM-Produktivität und der GFP-Fluoreszenz am Beispiel des Zellpools ZB1. Dieser Zellpool wurde aus der Transfektion mit dem Vektor pBIDG-sICAM erhalten, bei dem das therapeutische Protein sICAM und GFP gemeinsam von einer bicistronischen Transkriptionseinheit exprimiert werden. Der Pool wurde einem sequentiellen GFP-basierten FACS-Sortierung unterzogen. Nach jedem Sortierungsschritt (Sort) wurde die Konzentration des sICAMs im Zellkulturüberstand des Pools durch ELISA ermittelt und die spezifische Produktivität pro Zelle und Tag ($\text{pg}/\text{c}^*\text{d}$) berechnet. Jeder Datenpunkt stellt dabei den Durchschnitt aus mindestens drei Kultivierungspassagen dar.

5 Insgesamt wurden sechs Sorts durchgeführt.

Abbildung 5 zeigt die Isolierung von hochexprimierenden sICAM Zellen durch GFP-basierte FACS-Sortierung am Beispiel des Zellpools ZB1. Dieser Zellpool wurde aus der Transfektion mit dem Vektor pBIDG-sICAM erhalten, bei dem das therapeutische Protein sICAM und GFP gemeinsam von einer bicistronischen Transkriptionseinheit exprimiert werden. Der Pool wurde einer sequentiellen GFP-basierten FACS-Sortierung unterzogen. Nach jedem Sort wurde die Konzentration des sICAMs im Zellkulturüberstand des Pools durch ELISA ermittelt und die spezifische Produktivität pro Zelle und Tag ($\text{pg}/\text{c}^*\text{d}$) berechnet. Jeder Datenpunkt stellt dabei den Durchschnitt aus mindestens drei Kultivierungspassagen dar. Insgesamt wurden sechs Sorts durchgeführt.

Abbildung 6 zeigt die durch Kombination einer GFP-basierten Selektion mit einem MTX-Amplifikationsschritt erzielten Steigerungen in der sICAM-Produktivität am Beispiel des Zellpools ZB1. Dieser Zellpool, der aus der Transfektion mit dem Vektor pBIDG-sICAM erhalten wurde, wurde einer sequentiellen GFP-basierten FACS-Sortierung unterzogen. Nach dem vierten Sort bzw. sechsten Sort wurde eine DHFR-vermittelte Genamplifikation durch Zusatz von Methotrexat (MTX) zum Kultivierungsmedium durchgeführt (5 nM, 50 nM, 500 nM oder 2 μM MTX). Die Konzentration des sICAMs im Zellkulturüberstand des Pools wurde durch ELISA ermittelt und die spezifische Produktivität pro Zelle und Tag ($\text{pg}/\text{c}^*\text{d}$) berechnet. Jeder Datenpunkt stellt dabei den Durchschnitt aus mindestens drei Kultivierungspassagen dar.

Abbildung 7 zeigt den Viabilitätsverlauf von Zellpools nach Zugabe von unterschiedlichen hohen Dosen an Methotrexat zum Kultivierungsmedium. Der Zellpool ZB1, der aus der Transfektion mit dem Vektor pBIDG-sICAM (Abb.3) erhalten wurde, wurde einer sequentiellen GFP-basierten FACS-Sortierung unterzogen. Nach dem vierten Sort bzw. sechsten Sort wurde eine DHFR-vermittelte Genamplifikation durch Zusatz von Methotrexat (MTX) zum Kultivierungsmedium durchgeführt. Die Zellzahlen sowie die Viabilität wurden während der Selektionsphase durch Trypanblau-Färbung bestimmt und über mehrere Kultivierungstage (dic) hinweg verfolgt.



Abbildung 8 zeigt die Korrelation zwischen der Antikörper-Produktivität (mAk F19) und der GFP-Fluoreszenz am Beispiel des Zellpools ZB1. Dieser Zellpool wurde aus der Transfektion mit der Vektorkombination pBIDG-F19HC und pBIN-F19LC (Abb.3) erhalten. Der Pool wurde einer sequentiellen GFP-basierten FACS-Sortierung unterzogen. Nach jedem Sort wurde die Konzentration des Antikörpers F19 im Zellkulturüberstand des Pools durch ELISA ermittelt und die spezifische Produktivität pro Zelle und Tag ($\text{pg}/\text{c}^*\text{d}$) berechnet. Jeder Datenpunkt stellt dabei den Durchschnitt aus mindestens drei Kultivierungspassagen dar. Insgesamt wurden sechs Sorts durchgeführt.



Abbildung 9 zeigt die Isolierung von hochexprimierenden mAk F19 Zellpools durch eine GFP-basierte Selektion mittels FACS am Beispiel des Zellpools ZB1. Dieser Zellpool, der aus der Co-Transfektion mit den Vektoren pBIDG-F19HC und pBIN-F19LC (Abb.3) erhalten wurde, wurde einem sequentiellen GFP-basierten FACS-Sorting unterzogen. Die Konzentration des Antikörpers F19 im Zellkulturüberstand des Pools wurde nach jedem Sort durch ELISA ermittelt und die spezifische Produktivität pro Zelle und Tag ($\text{pg}/\text{c}^*\text{d}$) berechnet. Jeder Datenpunkt stellt dabei den Durchschnitt aus mindestens drei Kultivierungspassagen dar.

Detaillierte Beschreibung der Erfindung und bevorzugte Ausführungsformen

Der erfindungsgemäße Expressionsvektor enthält ein Gen, das für ein Protein von Interesse kodiert („Gen von Interesse“), in funktioneller Verknüpfung mit einem 5 Hamster-Ubiquitin/S27a-Promotor und einem Gen, das für ein fluoreszierendes Protein kodiert. Vorzugsweise enthält der Expressionsvektor noch ein amplifizierbares Selektionsmarkergen.

Hamster-Ubiquitin/S27a-Promotor

10

Der Ubiquitin/S27a-Promotor des Hamsters ist ein starker homologer Promotor, der in der WO 97/15664 beschrieben ist. Ein solcher Promotor weist vorzugsweise mindestens eines der folgenden Merkmale auf: GC-reicher Sequenzbereich, Sp1-Bindungsstelle, Polypyrimidinelement, Abwesenheit einer TATA-Box. Besonders 15 bevorzugt ist ein Promotor, der eine Sp1-Bindungsstelle bei Abwesenheit einer TATA-Box aufweist. Ferner ist ein solcher Promotor bevorzugt, der konstitutiv aktiviert ist und insbesondere unter serumhaltigen, serumarmen und serumfreien Zellkulturbedingungen gleichermaßen aktiv ist. In einer anderen Ausführungsform handelt es sich um einen induzierbaren Promotor, insbesondere um einen Promotor, 20 der durch Serumentzug aktiviert wird.

Eine besonders vorteilhafte Ausführungsform ist ein Promotor mit einer Nukleotidsequenz, die in Fig. 5 der WO 97/15664 enthalten ist. Besonders bevorzugt sind dabei Promotorsequenzen, in denen die Sequenz von Position -161 bis - 45 von 25 Fig. 5 enthalten ist.

Die in den Beispielen der vorliegenden Patentbeschreibung verwendeten Promotoren beinhalten jeweils ein DNA-Molekül mit der Sequenz von Position 1923 bis 2406 der SEQ ID NO: 1 des beiliegenden Sequenzprotokolls. Diese Sequenz 30 entspricht dem Fragment – 372 bis + 111 aus Fig. 5 der WO 97/15664 und repräsentiert den bevorzugten Promotor, d.h. ein bevorzugter Promotor sollte diesen Sequenzbereich umfassen. Ein anderer geeignetes Promotorfragment enthält die Sequenz von Position 2134 bis 2406 (entspricht - 161 bis + 111 in Fig. 5 der WO

97/15664). Ein Promotor, der lediglich die Sequenz von Position 2251 bis 2406 beinhaltet, ist nicht mehr funktionsfähig (entspricht Position - 45 bis + 111 in Fig. 5 der WO 97/15664). Eine Verlängerung der Promotorsequenz in 5'-Richtung ausgehend von Position 2134 ist möglich.

5

Es können auch funktionelle Subfragmente der vollständigen Hamster-Ubiquitin/S27a-Promotorsequenz sowie funktionelle Mutanten/Varianten der vollständigen Sequenz oder Subfragmente hiervon eingesetzt werden, die z.B. durch Substitutionen, Insertionen oder Deletionen modifiziert worden sind. Entsprechende 10 Subfragmente, Mutanten oder Varianten werden nachfolgend auch als „modifizierter Promotor“ bezeichnet.



Ein modifizierter Promotor, gegebenenfalls kombiniert mit weiteren regulatorischen Elementen, weist vorzugsweise eine Transkriptionsaktivität auf, die der des 15 Promotorfragments von Position 1923 bis 2406 der in SEQ ID NO:1 angegebenen Nukleotidsequenz (- 372 bis + 111 aus Fig. 5 der WO 97/15664) entspricht. Ein modifizierter Promotor erweist sich als tauglich im Sinne der Erfindung, wenn er über eine Transkriptionsaktivität verfügt, die mindestens 50%, besser mindestens 80%, noch besser mindestens 90%, und noch mehr bevorzugt mindestens 100% der 20 Aktivität des 1923 bis 2406 Fragments (- 372 bis + 111 Fragments) in einem vergleichenden Reportergen-Assay aufweist. Insbesondere bevorzugt sind 25 modifizierte Promotoren, die eine minimale Sequenzhomologie zur Wildtyp-Sequenz SEQ ID NO:1 des Hamster-Ubiquitin/S27a-Promotors von mindestens 80%, besser mindestens 85%, bevorzugt mindestens 90%, weiter bevorzugt mindestens 95% und besonders bevorzugt mindestens 97% aufweisen und über eine entsprechende Promotoraktivität in einem vergleichenden Reportergen-Assay verfügen.

In einem entsprechenden vergleichenden Reportergen-Assay werden die zu 30 testenden Promotorfragmente einschließlich der Referenzsequenz jeweils vor ein promotorloses Reportergen kloniert, das z.B. für Luciferase, sezernierte Alkalische Phosphatase oder grünes fluoreszierendes Protein (GFP) kodiert. Diese Konstrukte (Promotorsequenz + Reportergen) werden anschließend in die Testzellen, z.B. CHO-DG44, mittels Transfektion eingeführt und die Induktion der Reportergenexpression

durch das jeweilige Promotorfragment über die Bestimmung des Proteingehalts des Reportergens ermittelt. Ein entsprechender Test findet sich beispielhaft auch in Ausubel et al., Current Protocols in Molecular Biology, 1994, updated.

- 5 Die Promotorsequenz des Hamster-Ubiquitin/S27a-Promotors sowie die modifizierter Promotoren, die z.B. auch die 5'-untranslatierte Region bzw. ausgewählte Fragmente hiervon umfassen können, und die kodierende Region des Ubiquitin/S27a-Gens bzw. ausgewählte Fragmente hiervon, können von einem Fachmann in Kenntnis der in der WO 97/15664 beschriebenen Sequenz mit 10 verschiedenen Standardmethoden erhalten werden, wie z.B. in Sambrook et al., 1989; Ausubel et al., 1994 beschrieben. Ausgehend von der in WO 97/15664 beschriebenen Sequenz kann beispielsweise ein geeigneter Abschnitt ausgewählt und eine Oligonukleotid-Sonde chemisch synthetisiert werden, die die Sequenz dieses Abschnitts enthält. Mit einer solchen Sonde kann z.B. durch Hybridisierung 15 aus einer genomischen Bibliothek des Hamsters das Ubiquitin/S27a-Gen bzw. dessen 5'-untranslatierte Region oder sonstige Fragmente kloniert werden. Mittels des oben beschrieben Reportergen-Assays ist der Fachmann in der Lage, ohne erheblichen Aufwand promotoraktive Fragmente zu identifizieren und im Sinne der vorliegenden Erfindung zu verwenden. Die 5'-untranslatierte Region bzw. spezielle 20 Fragmente davon können auch leicht durch PCR-Amplifikation mit entsprechenden Primern aus genomischer DNA oder einer genomischen Bibliothek erhalten werden. Fragmente der 5'-untranslatierten Region können auch durch limitierten Exonuklease 25 III-Verdau aus größeren DNA-Fragmenten erhalten werden. Solche DNA-Moleküle können auch chemisch synthetisiert oder aus chemisch synthetisierten Fragmenten durch Ligation erzeugt werden.

Deletions-, Insertions- und Substitutionsmutanten lassen sich mittels „orts-spezifischer Mutagenese“ und/oder „PCR-basierten Mutagenese-Techniken“ erzeugen. Entsprechende Methoden sind beispielhaft in Lottspeich und Zorbas 30 (1998) (Kapitel 36.1), mit weiteren Verweisen, aufgeführt.

Es ist auch möglich über Kreuzhybridisierung mit Sonden aus dem 5'-untranslatierten Bereich des Hamster-Ubiquitin/S27a-Gens oder aus dem S27a-

Anteil des Hamster-Ubiquitin/S27a-Gens geeignete Promotorsequenzen aus korrespondierenden homologen Genen anderer, bevorzugt Säugerspezies, zu identifizieren und zu isolieren. Entsprechende Techniken sind beispielhaft in Lottspeich und Zorbas (1998) (Kapitel 23) beschrieben. „Homolog“ im Sinne der 5 Erfindung sind Gene, sofern ihre Nukleotidsequenz mindestens 70%, besser mindestens 80%, bevorzugt mindestens 90%, weiter bevorzugt mindestens 95% und besonders bevorzugt mindestens 97% Übereinstimmung mit der Nukleotidsequenz des Gens zeigt, zu dem es homolog ist.

10 Gen von Interesse

Das im erfindungsgemäßen Expressionsvektor enthaltene Gen von Interesse umfasst eine Nukleotidsequenz beliebiger Länge, die für ein Produkt von Interesse kodiert. Das Genprodukt oder auch „Produkt von Interesse“ ist in der Regel ein 15 Protein, Polypeptid, Peptid bzw. Fragment oder Derivat davon. Es kann aber auch RNA oder antisense RNA sein. Das Gen von Interesse kann in voller Länge, in verkürzter Form, als Fusionsgen oder markiertes Gen vorliegen. Es kann sich um genomische DNA oder vorzugsweise cDNA bzw. entsprechende Fragmente oder Fusionen handeln. Das Gen von Interesse kann die native Gensequenz darstellen, 20 mutiert oder auf sonstige Weise modifiziert sein. Derartige Modifikationen schließen Codon-Optimierungen zur Anpassung an eine bestimmte Wirtszelle und eine Humanisierung ein. Das Gen von Interesse kann z.B. für ein sekretiertes, zytoplasmatisches, kernlokalisiertes, membrangebundenes oder zelloberflächen-gebundenes Polypeptid kodieren.

25

Der Ausdruck „Nukleotidsequenz“ oder „Nukleinsäuresequenz“ bezeichnet ein Oligonukleotid, Nukleotide, Polynukleotide und deren Fragmente sowie DNA oder RNA genomischen oder synthetischen Ursprungs, die als Einzel- oder Doppelstrang vorliegen und den kodierenden oder den nicht-kodierenden Strang eines Gens 30 repräsentieren kann. Zur Modifikation von Nukleinsäuresequenzen können Standardtechniken, wie z.B. ortsspezifische Mutagenese oder PCR-vermittelte Mutagenese (z.B. in Sambrook et al., 1989 oder Ausubel et al., 1994 beschrieben), eingesetzt werden.

Unter „kodieren“ versteht man die Eigenschaft oder Fähigkeit einer spezifischen Sequenz von Nukleotiden in einer Nukleinsäure, beispielsweise einem Gen in einem Chromosom oder einer mRNA, als Matrize für die Synthese von anderen Polymeren und Makromolekülen wie z.B. rRNA, tRNA, mRNA, anderen RNA-Molekülen, cDNA oder Polypeptiden in einem biologischen Prozess zu dienen. Demnach kodiert ein Gen für ein Protein, wenn durch Transkription und nachfolgende Translation der mRNA das gewünschte Protein in einer Zelle oder einem anderen biologischen System produziert wird. Sowohl der kodierende Strang, dessen Nukleotidsequenz 5 identisch mit der mRNA-Sequenz ist und normalerweise auch in 10 Sequenzdatenbanken, z.B. EMBL oder GenBank, angegeben wird, als auch der als 15 Matrize für die Transkription dienende nicht-kodierende Strang eines Gens oder cDNA kann dabei als kodierend für ein Produkt oder Protein bezeichnet werden. Eine Nukleinsäure, die für ein Protein kodiert, schließt auch Nukleinsäuren mit ein, die auf 20 Grund des degenerierten genetischen Codes eine andere Nukleotidsequenzabfolge aufweisen, aber in der gleichen Aminosäuresequenz des Proteins resultieren. Nukleinsäuresequenzen, die für Proteine kodieren, können auch Introns enthalten.

Mit dem Ausdruck „cDNA“ werden Desoxyribonukleinsäuren bezeichnet, die durch 25 reverse Transkription und Synthese des zweiten DNA-Strangs aus einer von einem Gen produzierten mRNA oder anderen RNA hergestellt werden. Liegt die cDNA als doppelsträngiges DNA-Molekül vor, dann enthält sie sowohl einen kodierenden als 30 auch einen nicht-kodierenden Strang.

Mit dem Ausdruck „Intron“ werden nicht-kodierende Nukleotidsequenzen beliebiger Länge bezeichnet. Sie kommen natürlicherweise in vielen eukaryontischen Genen vor und werden von einem zuvor transkribierten mRNA-Vorläufer durch einen als Spleissen bezeichneten Prozess entfernt. Hierfür ist ein exaktes Herausschneiden des Introns am 5'- und 3'-Ende und eine korrekte Verbindung der entstehenden mRNA-Enden erforderlich, damit eine reife prozessierte mRNA mit dem für die erfolgreiche Proteinsynthese richtigen Leseraster hergestellt wird. Viele der an diesem Spleissing-Prozess beteiligten Spleiss-Donor- und Spleiss-Akzeptor-Stellen, 35 die sind die unmittelbar an den Exon-Intron- bzw. Intron-Exon-Grenzen vorliegenden

Sequenzen, sind mittlerweile charakterisiert worden. Für einen Überblick siehe Ohshima et al., 1987.

Protein von Interesse

5

Biopharmazeutisch bedeutsame Proteine/Polypeptide umfassen z.B. Antikörper, Enzyme, Cytokine, Lymphokine, Adhäsionsmoleküle, Rezeptoren sowie deren Derivate bzw. Fragmente, sind aber nicht auf diese beschränkt. Im allgemeinen sind alle Polypeptide bedeutsam, die als Agonisten oder Antagonisten wirken und/oder 10 therapeutische oder diagnostische Anwendung finden können.

- Der Ausdruck „Polypeptide“ wird für Aminosäuresequenzen oder Proteine verwendet und bezeichnet Polymere von Aminosäuren beliebiger Länge. Dieser Ausdruck schließt auch Proteine ein, die posttranslational durch Reaktionen wie beispielsweise 15 Glykosylierung, Phosphorylierung, Acetylierung oder Proteinprozessierung modifiziert werden. Die Struktur des Polypeptids kann z.B. durch Substitutionen, Deletionen oder Insertion von Aminosäuren, Fusion mit anderen Proteinen, unter Beibehaltung seiner biologischen Aktivität modifiziert werden.
- Beispiele für therapeutische Proteine sind Insulin, Insulin-ähnlicher Wachstumsfaktor, humanes Wachstumshormon (hGH) und andere Wachstumsfaktoren, Gewebeplasminogenaktivator (tPA), Erythropoetin (EPO), Cytokine, beispielsweise 20 Interleukine (IL) wie IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-11, IL-12, IL-13, IL-14, IL-15, IL-16, IL-17, IL-18, Interferon (IFN)-alpha, beta, gamma, omega 25 oder tau, Tumornekrosefaktor (TNF) wie z.B. TNF-alpha, beta oder gamma, TRAIL, G-CSF, GM-CSF, M-CSF, MCP-1 und VEGF. Weitere Beispiele sind monoklonale, polyklonale, multispezifische und einzelkettige (*single chain*) Antikörper und Fragmente davon, wie z.B. Fab, Fab', F(ab')₂, Fc und Fc'-Fragmente, leichte (L) und schwere (H) Immunglobulinketten und deren konstante, variable oder hypervariable 30 Regionen sowie Fv- und Fd-Fragmente (Chamov et al., 1999). Die Antikörper können humanen oder nicht-humanen Ursprungs sein. Auch humanisierte und chimäre Antikörper kommen in Frage.

Fab-Fragmente (Fragment antigen-binding = Fab) bestehen aus den variablen Regionen beider Ketten, die durch die angrenzenden konstanten Regionen zusammengehalten werden. Sie können z.B. durch Behandlung mit einer Protease, wie beispielsweise Papain, aus konventionellen Antikörpern erzeugt werden oder 5 aber auch durch DNA-Klonierung. Weitere Antikörperfragmente sind $F(ab')_2$ -Fragmente, die durch proteolytischen Verdau mit Pepsin hergestellt werden können.

Durch Genklonierung können auch verkürzte Antikörperfragmente hergestellt werden, die nur aus den variablen Region der schweren (VH) und der leichten Kette 10 (VL) bestehen. Diese werden als Fv-Fragmente (Fragment variable = Fragment des variablen Teils) bezeichnet. Da bei diesen Fv-Fragmenten die kovalente Verbindung über die Cysteinreste der konstanten Ketten nicht möglich ist, werden diese Fv-Fragmente oft anderweitig stabilisiert. Dazu werden die variablen Region der schweren und leichten Kette häufig mittels eines kurzen Peptidfragments von ca. 15 15 – 30 Aminosäuren, besonders bevorzugt 15 Aminosäuren, miteinander verknüpft. Auf diese Weise entsteht eine einzelne Polypeptidkette, in der VH und VL durch einen Peptidlinker miteinander verbunden sind. Solche Antikörperfragmente werden auch als *single-chain* Fv-Fragment (scFv) bezeichnet. Beispiele von scFv-Antikörpern sind bekannt und beschrieben, siehe z.B. Huston et al. (1988).

20 In den vergangenen Jahren wurden verschiedene Strategien entwickelt um multimer scFv-Derivate herzustellen. Die Intention besteht in der Erzeugung von 25 rekombinanten Antikörpern mit verbesserten pharmakokinetischen Eigenschaften und verstärkter Bindungsavidität. Zur Erreichung der Multimerisierung der scFv-Fragmente werden diese als Fusionsproteine mit Multimerisierungsdomänen hergestellt. Als Multimerisierungsdomänen können dabei z.B. die CH3-Region eines IgGs oder Helixstrukturen („coiled coil structure“) wie die Leucin-Zipper-Domänen fungieren. In anderen Strategien wird die Interaktion zwischen den VH- und VL-Regionen des scFv-Fragments für eine Multimerisierung genutzt (z.B. Dia-, Tri- und 30 Pentabodies).

Als „Diabody“ bezeichnet ein Fachmann ein bivalentes homodimeres scFv-Derivat. Die Verkürzung des Peptidlinkers im scFv-Moleküle auf 5 – 10 Aminosäuren

resultiert in der Bildung von Homodimeren durch Überlagerung von VH/VL-Ketten. Die Diabodies können zusätzlich durch eingeführte Disulfidbrücken stabilisiert werden. Beispiele von Diabodies finden sich in der Literatur, z.B. bei Perisic et al. (1994).

5

Als „Minibody“ bezeichnet der Fachmann ein bivalentes, homodimeres scFv-Derivat. Es besteht aus einem Fusionsprotein, das die CH3-Region eines Immunglobulins, vorzugsweise IgG, besonders bevorzugt IgG1, als Dimerisierungsregion enthält. Diese verbindet die scFv-Fragmente über eine Hinge-Region, ebenfalls von IgG, und 10 eine Linker-Region. Beispiele solcher Minibodies sind bei Hu et al. (1996) beschrieben.

15 Mit „Triabody“ bezeichnet der Fachmann ein trivalentes homotrimeres scFv-Derivat (Kortt et al., 1997). Die direkte Fusion von VH-VL ohne Verwendung einer Linkersequenz führt zur Ausbildung von Trimeren.

Bei den vom Fachmann als Mini-Antikörper bezeichneten Fragmenten, die eine bi-, tri- oder tetravalente Struktur haben, handelt es sich ebenfalls um Derivate von scFv-Fragmenten. Die Multimersierung wird dabei über di-, tri- oder tetramere „coiled coil“- 20 Strukturen erzielt (Pack et al., 1993 und 1995; Lovejoy et al., 1993).

Gen, das für ein fluoreszierendes Protein kodiert

Der erfindungsgemäße Expressionsvektor enthält ein für ein fluoreszierendes Protein 25 kodierendes Gen in funktioneller Verknüpfung mit dem Gen von Interesse und unter der Kontrolle des Hamster-Ubiquitin/S27a-Promotors, eines modifizierten Hamster-Ubiquitin/S27a-Promotors oder eines Homologen hiervon.

Bei dem fluoreszierenden Protein kann es sich z.B. um ein grün, blaugrün, blau, gelb 30 oder andersfarben fluoreszierendes Protein handeln. Ein spezielles Beispiel ist das grün fluoreszierende Protein (GFP) aus *Aequorea victoria* oder *Renilla reniformis* und daraus entwickelte Mutanten; siehe z.B. Bennet et al. (1998); Chalfie et al. (1994); WO 01/04306 und die dort zitierte Literatur.

Weitere fluoreszierende Proteine und dafür kodierende Gene sind in WO 00/34318, WO 00/34326, WO 00/34526 und WO 01/27150 beschrieben, die durch Bezugnahme hierin inkorporiert werden. Bei diesen fluoreszierenden Proteinen handelt es

5 sich um Fluorophore von nicht-biolumineszierenden Organismen der Spezies Anthozoa, beispielsweise von *Anemonia majano*, *Clavularia sp.*, *Zoanthus sp. I*, *Zoanthus sp. II*, *Discosoma striata*, *Discosoma sp. „red“*, *Discosoma sp. “green”*, *Discosoma sp. “Magenta”*, *Anemonia sulcata*.

10 Die erfindungsgemäß eingesetzten Fluoreszenzproteine beinhalten neben den Wildtyp-Proteinen auch natürliche oder gentechnologisch hergestellte Mutanten und -varianten, deren Fragmente, Derivate oder z.B. mit anderen Proteinen oder Peptiden fusionierte Varianten. Die eingebrachten Mutationen können dabei beispielsweise das Exzitations- oder Emissionsspektrum, die Chromophorenbildung,

15 den Extinktionskoeffizienten oder die Stabilität des Proteins verändern. Durch Codon-Optimierung kann zudem die Expression in Säugerzellen oder anderen Spezies verbessert werden. Erfindungsgemäß kann das fluoreszierende Protein auch in Fusion mit einem Selektionsmarker, bevorzugterweise mit einem amplifzierbaren Selektionsmarker wie beispielsweise der Dihydrofolat-Reduktase

20 (DHFR), eingesetzt werden.

25 Die von den fluoreszierenden Proteinen emittierte Fluoreszenz ermöglicht die Detektion der Proteine z.B. durch Durchflusszytometrie mit einem Fluoreszenz-aktivierten Zellsortierer (FACS) oder durch Fluoreszenzmikroskopie.

25

Weitere regulatorische Elemente

Der Hamster-Ubiquitin/S27a-Promotor kann zur Steigerung/Regulierung der Transkriptionsaktivität in einer Expressionskassette in funktionellen Zusammenhang

30 mit weiteren regulatorischen Sequenzen gebracht werden.

Beispielsweise kann der Promotor mit Enhancer-Sequenzen funktionell verknüpft werden, um die Transkriptionsaktivität zu steigern. Hierzu können ein oder mehrere

Enhancer und/oder mehrere Kopien einer Enhancer-Sequenz verwendet werden, beispielsweise ein CMV- oder SV40-Enhancer.

Mit dem Ausdruck „Enhancer“ wird eine Polynukleotidsequenz bezeichnet, die in *cis*-Lokalisierung auf die Aktivität eines Promotors einwirkt und so die Transkription eines mit diesem Promotor funktionell verknüpften Gens stimuliert. Im Gegensatz zu Promotoren ist die Wirkung der Enhancer positions- und orientierungsunabhängig und können somit vor oder hinter eine Transkriptionseinheit, innerhalb eines Introns oder selbst innerhalb der Kodierregion positioniert werden. Der Enhancer kann dabei sowohl in unmittelbarer Nähe der Transkriptionseinheit als auch in beträchtlichem Abstand zum Promotor lokalsiert sein. Auch eine physische und funktionelle Überlappung mit dem Promotor ist möglich. Dem Fachmann sind eine Vielzahl von Enhancern aus verschiedenen Quellen bekannt (und in Datenbanken wie GenBank hinterlegt, z.B. SV40 Enhancer, CMV Enhancer, Polyoma Enhancer, Adenovirus Enhancer) und als eigenständige oder innerhalb von Polynukleotidsequenzen klonierte Elemente verfügbar (z.B. bei ATCC hinterlegt oder aus kommerziellen und individuellen Quellen). Eine Vielzahl von Promotorsequenzen beinhalten auch Enhancersequenzen, wie z.B. der häufig verwendete CMV Promotor. Der humane CMV-Enhancer gehört dabei zu den stärksten bisher identifizierten Enhancern. Ein Beispiel für einen induzierbaren Enhancer ist der Metallothionein-Enhancer, der durch Glucocorticoide oder Schwermetalle stimuliert werden kann.

Eine weitere mögliche Modifikation ist z.B. die Einführung multipler Sp1-Bindungsstellen. Die Promotorsequenzen können ferner mit regulatorischen Sequenzen kombiniert werden, die eine Steuerung/Regulierung der Transkriptionsaktivität gestatten. So kann der Promotor reprimierbar/induzierbar gemacht werden. Dies kann beispielsweise durch die Verknüpfung mit Sequenzen geschehen, die Bindungsstellen für positiv oder negativ regulierende Transkriptionsfaktoren darstellen. Der oben genannte Transkriptionsfaktor SP-1 beispielsweise hat einen positiven Einfluß auf die Transkriptionsaktivität. Ein weiteres Beispiel ist die Bindungsstelle für das Aktivatorprotein AP-1, das sowohl in positiver als auch in negativer Weise auf die Transkription einwirken kann. Die Aktivität des AP-1 kann durch verschiedenste Faktoren, wie z.B. Wachstumsfaktoren, Zytokine und Serum,

gesteuert werden (Faisst et al., 1992, und Referenzen darin). Die Transkriptions-
effizienz kann auch dadurch gesteigert werden, dass die Promotorsequenz durch
Mutation (Substitution, Insertion oder Deletion) von einer, zwei, drei oder mehr Basen
verändert wird und dann in einem Reportergen-Assay bestimmt wird, ob sich
5 dadurch die Promotoraktivität erhöht.

Grundsätzlich umfassen die zusätzlichen regulatorischen Elemente andere
Promotoren als den Hamster-Ubiquitin/S27a-Promotor, Enhancer, Terminations- und
Polyadenylierungssignale und weitere Expressionskontrollelemente. Für die
10 verschiedenen Zelltypen sind sowohl induzierbare als auch konstitutive
regulatorische Sequenzen bekannt. „Transkriptionsregulatorische Elemente“
umfassen gewöhnlich einen Promotor stromaufwärts von der zu exprimierenden
Gensequenz, Transkriptionsinitiations- und –terminationsstellen sowie ein
Polyadenylierungssignal.

15 Als „Promotor“ wird eine Polynukleotidsequenz bezeichnet, die die Transkription der
mit ihr funktionell verknüpften Gene oder Sequenzen ermöglicht und kontrolliert. Ein
Promotor enthält Erkennungssequenzen für die Bindung der RNA-Polymerase und
die Initiationsstelle der Transkription (Transkriptionsinitiationsstelle). Zur Expression
20 einer gewünschten Sequenz in einem bestimmten Zelltyp oder einer Wirtszelle muss
jeweils ein geeigneter, funktionaler Promotor gewählt werden. Der Fachmann kennt
eine Vielzahl von Promotoren aus verschiedenen Quellen, einschließlich
25 konstitutiver, induzierbarer und reprimierbarer Promotoren. Sie sind in Datenbanken,
z.B. GenBank, hinterlegt und können als eigenständige oder innerhalb von
Polynukleotidsequenzen klonierte Elemente von kommerziellen oder individuellen
Quellen bezogen werden. In induzierbaren Promotoren kann die Aktivität des
Promotors in Reaktion auf ein Signal reduziert oder verstärkt werden. Ein Beispiel für
30 einen induzierbaren Promotor stellt der Tetracyclin (tet)-Promotor dar. Dieser enthält
Tetracyclin-Operatorsequenzen (tetO), die durch ein Tetracyclin-reguliertes
Transaktivatorprotein (tTA) induziert werden können. In der Anwesenheit von
Tetracyclin wird die Bindung von tTA an tetO inhibiert. Beispiele für weitere
induzierbare Promotoren sind der jun-, fos-, Metallothionin- und Hitzeschockpromotor
(siehe auch Sambrook et al., 1989; Gossen et al., 1994). Unter den Promotoren, die

für eine hohe Expression in Eukaryonten besonders gut geeignet sind, befinden sich der SV40 early Promotor, der Adenovirus major late Promotor, der Maus Metallothionin-I Promotor, die lange terminale Repeatregion des Rous Sarcoma Virus und der early Promotor des humanen Cytomegalie-Virus. Beispiele für andere 5 heterologe Säugerpromotoren sind der/die Aktin-, Immunglobulin-, oder Hitzeschockpromotor(en).

Der Ausdruck „Transkriptionsinitiationsstelle“ bezieht sich auf eine Nukleinsäure in dem Konstrukt, die der ersten Nukleinsäure entspricht, welche in das primäre 10 Transkript, d.h. den mRNA-Precursor, inkorporiert wird. Die Transkriptionsinitiationsstelle kann mit den Promotorsequenzen überlappen.

Der Ausdruck „Transkriptionsterminationsstelle“ bezieht sich auf eine Nukleotidsequenz, die normalerweise am 3'-Ende des Gens von Interesse oder des 15 zu transkribierenden Genabschnitts vorhanden ist und die den Abbruch der Transkription durch RNA-Polymerase bewirkt.

Das „Polyadenylierungssignal“ ist eine Signalsequenz, welche die Spaltung an einer spezifischen Stelle am 3'-Ende der eukaryontischen mRNA und den posttranskriptionellen Einbau einer Sequenz von etwa 100-200 Adeninnukleotiden (polyA-Schwanz) am gespaltenen 3'-Ende verursacht. Das Polyadenylierungssignal umfasst 20 die Sequenz AATAAA etwa 10-30 Nukleotide stromaufwärts von der Spaltstelle sowie eine stromabwärts gelegene Sequenz. Es sind verschiedene Polyadenylierungselemente bekannt, z.B. tk polyA, SV40 late und early polyA oder BGH 25 polyA (z.B. beschrieben in US 5,122,458).

In einer bevorzugten Ausführungsform der vorliegenden Erfindung verfügt jede Transkriptionseinheit über einen Promotor oder ein Promotor/Enhancer-Element, ein Gen von Interesse und/oder ein Markergen, sowie über ein 30 Transkriptionterminationselement. In einer weiter bevorzugten Ausführungsform enthält die Transkriptionseinheit weitere translationsregulatorische Einheiten.

„Translationsregulatorische Elemente“ umfassen eine Translationsinitiationsstelle (AUG), ein Stoppcodon und ein polyA-Signal für jedes zu exprimierende Polypeptid. Für eine optimale Expression kann es günstig sein, 5'- und/oder 3'-nichttranslatierte Bereiche der zu exprimierenden Nukleinsäuresequenz zu entfernen, hinzuzufügen 5 oder zu ändern, um potentielle zusätzliche ungeeignete Translationsinitiationscodons oder andere Sequenzen, welche die Expression auf dem Niveau der Transkription oder Expression beeinträchtigen können, zu eliminieren. Um die Expression zu fördern, kann man alternativ ribosomale Konsensus-Bindungsstellen unmittelbar stromaufwärts vom Startcodon insertieren. Um ein sekretiertes Polypeptid zu 10 produzieren, enthält das Gen von Interesse gewöhnlich eine Signalsequenz, welche ~~die~~ für ein Signalvorläuferpeptid kodiert, die das synthetisierte Polypeptid zu und durch die ER-Membran transportiert. Die Signalsequenz befindet sich oft, jedoch nicht immer, am Aminoterminus des sekretierten Proteins und wird durch Signal-Peptidasen abgespalten, nachdem das Protein durch die ER-Membran geschleust 15 wurde. Die Gensequenz wird gewöhnlich, jedoch nicht notwendigerweise, eine eigene Signalsequenz enthalten. Wenn die native Signalsequenz nicht vorhanden ist, kann in bekannter Weise eine heterologe Signalsequenz eingeführt werden. Dem Fachmann sind zahlreiche solcher Signalsequenzen bekannt und in Sequenzdatenbanken wie GenBank und EMBL hinterlegt.

20 Ein erfindungsgemäß besonderes bedeutsames regulatorisches Element ist die ~~die~~ interne Ribosomenbindungsstelle (IRES). Das „IRES-Element“ umfasst eine Sequenz, welche die Translationsinitiation unabhängig von einer 5'-terminalen Methylguanosiniumkappe (CAP-Struktur) sowie dem stromaufwärts gelegenen Gen 25 funktionell bewerkstelligt und in einer tierischen Zelle die Translation zweier Cistrone (offener Leseraster) von einem einzigen Transkript ermöglicht. Das IRES-Element stellt eine unabhängige Ribosomenbindungsstelle für die Translation des unmittelbar stromabwärts gelegenen offenen Leserasters zur Verfügung. Im Gegensatz zu bakterieller mRNA, die multicistronisch sein kann, d.h. für mehrere verschiedene 30 Polypeptide oder Produkte kodieren kann, die nacheinander von der mRNA translatiert werden, sind die meisten mRNAs von Tierzellen monocistronisch und kodieren nur für ein einziges Protein oder Produkt. Bei einem multicistronischen Transkript in einer eukaryontischen Zelle würde die Translation von der

stromaufwärts am nächsten gelegenen Translationsinitiationsstelle initiiert und vom ersten Stoppcodon beendet werden, worauf das Transkript aus dem Ribosom freigesetzt würde. Bei der Translation entstünde somit nur das erste von der mRNA kodierte Polypeptid oder Produkt. Dagegen ermöglicht ein multicistronisches 5 Transkript mit einem IRES-Element, das mit dem zweiten oder weiteren offenen Leserastern in dem Transkript funktionell verknüpft ist, die anschließende Translation des stromabwärts gelegenen offenen Leserasters, so dass in der eukaryontischen Zelle zwei oder mehr, von demselben Transkript kodierte Polypeptide oder Produkte produziert werden.

10

Das IRES-Element kann von verschiedener Länge und von verschiedenem Ursprung 10 sein und z.B. vom Encephalomyocarditisvirus (EMCV) oder anderen Picornaviren stammen. In der Literatur sind verschiedene IRES-Sequenzen und deren Anwendung bei der Konstruktion von Vektoren beschrieben worden; siehe z.B. 15 Pelletier et al., 1988; Jang et al., 1989; Davies et al., 1992; Adam et al., 1991; Morgan et al., 1992; Sugimoto et al., 1994; Ramesh et al., 1996, Mosser et al., 1997.

Die stromabwärts gelegene Gensequenz ist mit dem 3'-Ende des IRES-Elements 20 funktionell verknüpft, d.h. der Abstand wird so gewählt, dass die Expression des Gens nicht oder nur marginal beeinflusst wird bzw. eine für den Zweck ausreichende Expression aufweist. Der optimale und zulässige Abstand zwischen dem IRES- 25 Element und dem Startcodon des stromabwärts gelegenen Gens für eine noch ausreichende Expression lässt sich in einfachen Versuchen durch Variation des Abstands und Bestimmung der Expressionsrate als Funktion des Abstandes mit Hilfe von Reportergen-Assays ermitteln.

Durch die geschilderten Maßnahmen kann eine optimierte Expressionskassette erhalten werden, die von hohem Nutzen für die Expression heterologer Genprodukte ist. Eine durch eine oder mehrere solcher Maßnahmen erhaltene Expressions- 30 kassette ist daher ebenfalls Gegenstand der Erfindung.

Amplifizierbares Selektionsmarkergen

Ein bevorzugter erfindungsgemäßer Vektor enthält zusätzlich ein amplifizierbares Selektionsmarkergen, das eine Amplifikation des amplifizierbaren Markergens und vorzugsweise die Co-Amplifikation einer Transkriptionseinheit, bestehend aus dem Hamster-Ubiquitin/S27a-Gens, dem Gen von Interesse und dem Gen für das fluoreszierende Protein, ermöglicht. Hierzu werden die mit einem entsprechenden Expressionsvektor transfizierten Wirtszellen in Gegenwart eines geeigneten Selektionsmittels kultiviert, so dass sich nur solche Wirtszellen vermehren (lassen), die über mehrere Genkopien zumindest des amplifizierbaren Selektionsmarkergens verfügen. Vorzugsweise wird dies durch eine stufenweise Kultivierung der Zellen in Gegenwart steigender Mengen an Selektionsmittel erreicht.

Das amplifizierbare Selektionsmarkergen kodiert gewöhnlich für ein Enzym, das für das Wachstum von eukaryontischen Zellen unter bestimmten Kultivierungsbedingungen erforderlich ist. Beispielsweise kann das amplifizierbare Selektionsmarkergen für Dihydrofolat-Reduktase (DHFR) kodieren. In diesem Fall wird das Gen amplifiziert, wenn man eine damit transfizierte Wirtszelle in Gegenwart des Selektionsmittels Methotrexat (MTX) kultiviert.

In der folgenden Tabelle 1 sind Beispiele für weitere erfindungsgemäß verwendbare amplifizierbare Selektionsmarkergene und die dazugehörigen Selektionsmittel angegeben, die in einem Überblick bei Kaufman, Methods in Enzymology, 185:537-566 (1990) beschrieben sind.

25

Tabelle 1: Amplifizierbare Selektionsmarkergene

Amplifizierbares Selektionsmarkergen	Hinterlegungsnummer	Selektionsmittel
Dihydrofolat-Reduktase	M19869 (Hamster) E00236 (Maus)	Methotrexat (MTX)
Metallothionein	D10551 (Hamster) M13003 (Human) M11794 (Ratte)	Cadmium
CAD (Carbamoylphosphat-Synthetase : Aspartat-	M23652 (Hamster) D78586 (Human)	N-Phosphoacetyl-L-aspartat

Transcarbamylase: Dihydroorotase)		
Adenosin-Deaminase	K02567 (Human) M10319 (Maus)	Xyl-A- oder Adenosin, 2'Deoxycoformycin
AMP (Adenylat)- Deaminase	D12775 (Human) J02811 (Ratte)	Adenin, Azaserin, Coformycin
UMP-Synthase	J03626 (Human)	6-Azauridin, Pyrazofuran
IMP 5'-Dehydrogenase	J04209 (Hamster) J04208 (Human) M33934 (Maus)	Mycophenolsäure
Xanthin-Guanin- Phosphoribosyltransferase	X00221 (E. coli)	Mycophenolsäure mit limitierendem Xanthin
Mutanten-HGPRTase oder Mutanten-Thymidin-Kinase	J00060 (Hamster) M13542, K02581 (Human) J00423, M68489 (Maus) M63983 (Ratte) M36160 (Herpes virus)	Hypoxanthin, Aminopterin, und Thymidin (HAT)
Thymidylat-Synthetase	D00596 (Human) M13019 (Maus) L12138 (Ratte)	5-Fluorodeoxyuridin
P-Glycoprotein 170 (MDR1)	AF016535 (Human) J03398 (Maus)	mehrere Arzneistoffe, z.B. Adriamycin, Vincristin, Colchicin
Ribonukleotid-Reduktase	M124223, K02927 (Maus)	Aphidicolin
Glutamin-Synthetase	AF150961 (Hamster) U09114, M60803 (Maus) M29579 (Ratte)	Methioninsulfoximin (MSX)
Asparagin-Synthetase	M27838 (Hamster) M27396 (Human) U38940 (Maus) U07202 (Ratte)	β -Aspartylhydroxamat, Albizzin, 5'Azacytidin
Argininosuccinat- Synthetase	X01630 (Human) M31690 (Maus) M26198 (Rind)	Canavanin
Ornithin-Decarboxylase	M34158 (Human) J03733 (Maus) M16982 (Ratte)	α -Difluormethylornithin
HMG-CoA-Reduktase	L00183, M12705 (Hamster) M11058 (Human)	Compactin
N-Acetylglucosaminyl- Transferase	M55621 (Human)	Tunicamycin
Threonyl-tRNA-Synthetase	M63180 (Human)	Borrelidin
Na ⁺ K ⁺ -ATPase	J05096 (Human) M14511 (Ratte)	Ouabain

Als amplifizierbares Selektionsmarkergen wird erfindungsgemäß ein Gen bevorzugt, das für ein Polypeptid mit der Funktion von DHFR kodiert, z.B. für DHFR oder ein

Fusionsprotein aus dem fluoreszierenden Protein und DHFR. DHFR ist für die Biosynthese von Purinen erforderlich. Zellen, denen das DHFR-Gen fehlt, können in Purin-defizientem Medium nicht wachsen. Das DHFR-Gen ist deshalb ein nützlicher Selektionsmarker zur Selektion und Amplifikation von Genen in Zellen, die in Purin-freiem Medium kultiviert werden. Das Selektionsmittel, das in Verbindung mit dem DHFR-Gen verwendet wird, ist Methotrexat (MTX). Die vorliegende Erfindung schließt deshalb eine Methode zur Herstellung von hochproduzierenden rekombinanten Wirtszellen ein, wobei sie folgende Schritte enthält: (i) Transfektion der Wirtszellen mit Genen, die zumindest für ein Protein von Interesse, ein fluoreszierendes Protein und DHFR kodieren, (ii) Kultivierung der Zellen unter Bedingungen, die eine Expression der verschiedenen Gene ermöglichen, und (iii) die Amplifikation dieser co-integrierten Gene durch Kultivierung der Zellen in Gegenwart eines Selektionsmittels, das die Amplifikation zumindest des amplifizierbaren selektionierbaren Markergens erlaubt, wie z.B. Methotrexat. Bevorzugt werden die transfizierten Zellen hierbei in Hypoxanthin/Thymidin-freiem Medium in der Abwesenheit von Serum und unter Zugabe steigender Konzentrationen an MTX kultiviert. Verzugsweise beträgt die Konzentration an MTX bei dem ersten Amplifikationsschritt zumindest 200 nM, in einer noch mehr bevorzugten Ausführungsform zumindest 500 nM und kann stufenweise auf bis zu 1 μ M gesteigert werden. Im Einzelfall können auch Konzentration von über 1 μ M verwendet werden.

Säugerzellen, vorzugsweise Mausmyeloma- und Hamsterzellen, sind bevorzugte Wirtszellen für den Einsatz von DHFR als amplifizierbaren Selektionsmarker. Besonders bevorzugt sind die Zelllinien CHO-DUKX (ATCC CRL-9096) und CHO-DG44 (Urlaub et al., 1983), da sie bedingt durch Mutation keine eigene DHFR-Aktivität aufweisen. Um die DHFR-bedingte Amplifikation auch in anderen Zelltypen anwenden zu können, die über eine eigene endogene DHFR-Aktivität verfügen, kann bei der Transfektion ein mutiertes DHFR-Gen verwendet werden, das für ein Protein mit einer reduzierten Sensitivität gegenüber Methotrexat kodiert (Simonson et al., 1983; Wigler et al., 1980; Haber et al., 1982).

Herstellung erfindungsgemäßer Expressionsvektoren

Die Herstellung des erfindungsgemäßen Expressionsvektors kann grundsätzlich nach herkömmlichen, dem Fachmann geläufigen Methoden, wie z.B. bei Sambrook et al. (1989), beschrieben, erfolgen. Dort findet sich auch eine Beschreibung der funktionellen Komponenten eines Vektors, z.B. geeigneter Promotoren (zusätzlich zum Hamster-Ubiquitin/S27a-Promotor), Enhancer, Terminations- und Polyadenylierungssignale, Antibiotikaresistenzgene, Selektionsmarker, Replikationsstartpunkte und Spleisssignale. Zur Herstellung können herkömmliche Klonierungsvektoren verwendet werden, z.B. Plasmide, Bacteriophagen, Phagemide, Cosmide oder virale Vektoren wie Baculovirus, Retroviren, Adenoviren, Adeno-assoziierte Viren und Herpes simplex-Virus, aber auch künstliche Chromosomen/Minichromosomen. Die eukaryontischen Expressionsvektoren enthalten typischerweise auch prokaryontische Sequenzen wie z.B. Replikationsursprung und Antibiotikaresistenzgene, die die Vermehrung und Selektion des Vektors in Bakterien ermöglichen. Eine Vielzahl von eukaryontischen Expressionsvektoren, die multiple Klonierungsstellen zur Einführung einer Polynukleotidsequenz enthalten, sind bekannt und einige sind kommerziell bei verschiedenen Firmen wie Stratagene, La Jolla, CA, USA; Invitrogen, Carlsbad, CA, USA; Promega, Madison, WI, USA oder BD Biosciences Clontech, Palo Alto, CA, USA erhältlich.

Auf eine dem Fachmann geläufige Weise werden der Hamster-Ubiquitin/S27a-Promotor, das Gen von Interesse, das für ein fluoreszierendes Protein kodierende Gen, vorzugsweise auch das amplifizierbare Selektionsmarkergen, z.B. Dihydrofolat-Reduktase, sowie gegebenenfalls zusätzliche regulatorische Elemente wie die interne Ribosomenbindungsstelle (IRES), Enhancer oder ein Polyadenylierungssignal in den Expressionsvektor eingeführt. Ein erfindungsgemäßer Expressionsvektor enthält minimal einen Ubiquitin/S27a-Promotor, das Gen von Interesse, und das für ein fluoreszierendes Protein kodierende Gen. Vorzugsweise enthält der Expressionsvektor noch ein amplifizierbares Selektionsmarkergen. Erfindungsgemäß ist hierbei auch die Verwendung modifizierter Ubiquitin/S27a-Promotoren, z.B. wie die in der vorliegenden Anmeldung beschriebenen modifizierten Ubiquitin/S27a-Promotoren. Besonders bevorzugt ist ein Expressionsvektor, bei dem der Ubiquitin-Promotor, das

Gen von Interesse und das Gen, welches für ein fluoreszierendes Protein kodiert, funktionell miteinander verknüpft sind oder in funktioneller Verknüpfung stehen.

Im Rahmen der vorliegenden Beschreibung bezieht sich der Ausdruck „funktionelle Verknüpfung“ bzw. „funktionell verknüpft“ auf zwei oder mehr Nukleinsäuresequenzen oder –teilsequenzen, die so positioniert sind, dass sie ihre beabsichtigte Funktion ausüben können. Beispielsweise ist ein Promotor/Enhancer funktionell mit einer kodierenden Gensequenz funktionell verknüpft, wenn er in cis-Stellung die Transkription der verknüpften Gensequenz kontrollieren oder modulieren kann. Im Allgemeinen, jedoch nicht notwendigerweise, befinden sich funktionell verknüpfte DNA-Sequenzen in enger Nachbarschaft und, sofern zwei kodierende Gensequenzen verknüpft werden oder im Falle einer Sekretionssignalsequenz, im gleichen Leseraster. Obwohl sich ein funktionell verknüpfter Promotor im Allgemeinen stromaufwärts von der kodierenden Gensequenz befindet, muss er nicht notwendigerweise eng benachbart sein. Enhancer müssen ebenfalls nicht in enger Nachbarschaft vorliegen, solange sie die Transkription der Gensequenz begünstigen. Zu diesem Zweck können sie sowohl stromaufwärts als auch stromabwärts von der Gensequenz vorliegen, gegebenenfalls in einem Abstand. Eine Polyadenylierungsstelle ist funktionell mit einer Gensequenz verknüpft, wenn sie am 3'-Ende der Gensequenz derart positioniert ist, dass die Transkription über die kodierende Sequenz bis hin zum Polyadenylierungssignal fortschreitet. Die Verknüpfung kann nach üblichen rekombinanten Methoden erfolgen, z.B. mittels der PCR-Technik, durch Ligation an geeigneten Restriktionsschnittstellen oder durch Spleissen. Wenn keine geeigneten Restriktionsschnittstellen vorhanden sind, können in an sich bekannter Weise synthetische Oligonukleotid-Linker oder Adaptoren verwendet werden. Erfindungsgemäß erfolgt die funktionelle Verknüpfung vorzugsweise nicht über Intronsequenzen.

In einer der beschriebenen Ausführungsformen sind Ubiquitin/S27a-Promotor bzw. eine modifizierte Form hiervon, das Gen von Interesse und das für ein fluoreszierendes Protein kodierende Gen funktionell mit einander verknüpft. Dies meint z.B. das sowohl das Gen von Interesse als auch das für ein fluoreszierendes Protein kodierende Gen ausgehend von dem selben Ubiquitin/S27a-Promotor bzw. einer

modifizierten Form hiervon, exprimiert werden. In einer besonders bevorzugten Ausführungsform erfolgt die funktionelle Verknüpfung über ein IRES-Element, so dass von beiden Genen eine bicistronische mRNA synthetisiert wird. Der erfindungsgemäße Expressionsvektor kann zusätzlich Enhancer-Elemente enthalten, 5 die funktionell auf einen oder mehrere Promotoren wirken. Besonders bevorzugt ist ein Expressionsvektor, bei dem der Ubiquitin/S27a-Promotor bzw. eine modifizierte Form hiervon mit einem Enhancer-Element, z.B. einen SV40-Enhancer oder einem CMV-Enhancer-Element verknüpft ist.

10 Grundsätzlich kann die Expression der Gene innerhalb eines Expressionsvektors von einer oder mehreren Transkriptionseinheiten aus erfolgen. Als „Transkriptionseinheit“ wird dabei eine Region definiert, die ein oder mehr zu transkribierende Gene enthält. Dabei sind die Gene innerhalb einer Transkriptionseinheit funktionell derart miteinander verknüpft, dass alle Gene innerhalb einer solchen Einheit unter der 15 transkriptionellen Kontrolle desselben Promotors oder Promotors/Enhancers stehen. Als Resultat dieser transkriptionellen Verknüpfung von Genen kann mehr als ein Protein oder Produkt von einer Transkriptionseinheit aus transkribiert und somit exprimiert werden. Jede Transkriptionseinheit enthält dabei die regulatorischen Elemente, die für die Transkription und die Translation der in ihr enthaltenen 20 Gensequenzen erforderlich sind. Jede Transkriptionseinheit kann dabei die gleichen oder verschiedene regulatorische Elemente enthalten. Für die funktionelle 25 Verknüpfung der Gene innerhalb einer Transkriptionseinheit können IRES-Elemente oder Introns verwendet werden.

25 Der Expressionsvektor kann eine einzige Transkriptionseinheit zur Expression des Gens von Interesse, des Gens für das Fluoreszenzprotein und des amplifizierbaren Selektionsmarker enthalten. Alternativ können diese Gene auch in zwei oder mehr Transkriptionseinheiten angeordnet sein. Dabei sind verschiedene Kombinationen der Gene innerhalb einer Transkriptionseinheit möglich. In einer weiteren 30 Ausführungsform der vorliegenden Erfindung kann mehr als ein Expressionsvektor, bestehend aus ein, zwei oder mehr Transkriptionseinheiten, in eine Wirtszelle durch Co-Transfektion oder in aufeinanderfolgenden Transfektionen in beliebiger Reihenfolge eingeführt werden. Jede Kombination von regulatorischen Elementen

und Genen auf jedem Vektor kann gewählt werden, so lange eine ausreichende Expression der Transkriptionseinheiten gewährleistet ist. Falls erforderlich können weitere regulatorische Elemente und Gene, wie z.B. zusätzliche Gene von Interesse oder Selektionsmarker, auf den Expressionsvektoren positioniert werden.

5

Demnach kann der erfindungsgemäße Expressionsvektor das Gen, das für ein fluoreszierendes Protein kodiert, und das amplifizierbare Selektionsmarkergen in einer oder in zwei separaten Transkriptionseinheiten enthalten. Jede Transkriptionseinheit kann ein oder mehrere Genprodukte transkribieren und 10 exprimieren. Wenn beide Gene in einer Transkriptionseinheit enthalten sind, stehen sie unter der Kontrolle desselben Promotors oder Promotors/Enhancers, wobei vorzugsweise ein IRES-Element verwendet wird, um die funktionelle Verknüpfung aller Komponenten zu gewährleisten. Wenn das Gen, das für ein fluoreszierendes Protein kodiert, und das amplifizierbare Selektionsmarkergen in zwei separaten 15 Transkriptionseinheiten enthalten sind, können sie unter der Kontrolle desselben oder verschiedener Promotoren/Enhancer stehen. Vorzugsweise verwendet man jedoch für das Selektionsmarkergen seinen natürlichen oder einen schwächeren heterologen Promotor, z.B. den SV40 early Promotor, und setzt vorzugsweise auch keinen Enhancer ein. Expressionsvektoren mit zwei separaten 20 Transkriptionseinheiten sind im Rahmen der Erfindung bevorzugt. Hierbei enthält die eine (bicistronische) Transkriptionseinheit das Gen von Interesse und das für ein fluoreszierendes Protein kodierende Gen, während die andere Transkriptionseinheit das amplifizierbare Selektionsmarkergen enthält. Vorzugsweise ist jede 25 Transkriptionseinheit am 3'-Ende durch eine Sequenz begrenzt, die für ein polyA-Signal, vorzugsweise tk polyA, BGH polyA oder SV40 polyA kodiert.

Erfindungsgemäß sind auch solche Vektoren, die anstelle des Gens von Interesse lediglich eine multiple Klonierungsstelle aufweisen, die die Klonierung des Gens von Interesse über Erkennungssequenzen für Restriktionsendonukleasen ermöglicht. Im 30 Stand der Technik sind zahlreiche Erkennungssequenzen für die verschiedensten Restriktionsendonukleasen, sowie die hierzu gehörigen Restriktionsendonukleasen bekannt. Bevorzugt werden Sequenzen verwendet, die aus zumindest 6 Nukleotiden

als Erkennungssequenz bestehen. Eine Auflistung geeigneter Erkennungssequenzen finden sich beispielsweise in Sambrook et al. (1989).

Wirtszellen

5

Zur Transfektion mit dem erfindungsgemäßen Expressionsvektor werden eukaryontische Wirtszellen verwendet, vorzugsweise Säugerzellen und insbesondere Nagerzellen wie z.B. Mäuse-, Ratten- und Hamster-Zelllinien. Die erfolgreiche Transfektion der entsprechender Zellen mit einem erfindungsgemäßen Expressionsvektor resultiert in transformierten, genetisch modifizierten, rekombinanten oder transgenen Zellen, die ebenfalls Gegenstand der vorliegenden Erfindung sind.

Im Rahmen der Erfindung bevorzugte Wirtszellen sind Hamsterzellen wie z.B. BHK21, BHK TK⁻, CHO, CHO-K1, CHO-DUKX, CHO-DUKX B1 und CHO-DG44 Zellen oder Derivate/Abkömmlinge dieser Zelllinien. Besonders bevorzugt sind CHO-DG44, CHO-DUKX, CHO-K1 und BHK21 Zellen, insbesondere CHO-DG44 und CHO-DUKX Zellen. Ebenfalls geeignet sind Myelomzellen der Maus, vorzugsweise NS0 und Sp2/0 Zellen sowie Derivate/Abkömmlinge dieser Zelllinien.

20

Beispiele für Hamster- und Mäusezellen, die erfindungsgemäß angewandt werden können, sind in der folgenden Tabelle 2 angegeben. Aber auch Derivate und Abkömmlinge dieser Zellen, andere Säugerzellen, einschließlich aber nicht beschränkt auf Zelllinien von Mensch, Maus, Ratte, Affen, Nagetieren, oder eukaryontische Zellen, einschließlich aber nicht beschränkt auf Hefe-, Insekten- und Pflanzenzellen, können ebenfalls als Wirtszellen zur Produktion von biopharmazeutischen Proteinen verwendet werden.

Tabelle 2: Hamster- und Mäuse-Produktionszelllinien

Zelllinie	Hinterlegungsnummer
NS0	ECACC No. 85110503
Sp2/0-Ag14	ATCC CRL-1581
BHK21	ATCC CCL-10
BHK TK ⁻	ECACC No. 85011423

HaK	ATCC CCL-15
2254-62.2 (BHK-21-Derivat)	ATCC CRL-8544
CHO	ECACC No. 8505302
CHO-K1	ATCC CCL-61
CHO-DUKX (= CHO duk ⁻ , CHO/dhfr ⁻)	ATCC CRL-9096
CHO-DUKX B1	ATCC CRL-9010
CHO-DG44	Urlaub et al., Cell 33[2], 405-412, 1983
CHO Pro-5	ATCC CRL-1781
V79	ATCC CCC-93
B14AF28-G3	ATCC CCL-14
CHL	ECACC No. 87111906

Die Transfektion der eukaryontischen Wirtszellen mit einem Polynukleotid oder einem der erfindungsgemäßen Expressionsvektor erfolgt nach üblichen Methoden (Sambrook et al., 1989; Ausubel et al., 1994). Geeignete Transfektionsmethoden sind z.B. die Liposomen-vermittelte Transfektion, Calciumphosphat-Copräzipitation, Elektroporation, Polykationen (z.B. DEAE-Dextran)-vermittelte Transfektion, Protoplastenfusion, Mikroinjektion und virale Infektionen. Erfindungsgemäß wird vorzugsweise eine stabile Transfektion durchgeführt, wobei die Konstrukte entweder in das Genom der Wirtszelle oder ein artifizielles Chromosom/Minichromosom integriert werden oder in stabiler Weise episomal in der Wirtszelle enthalten sind. Die Transfektionsmethode, die die optimale Transfektionsfrequenz und Expression des heterologen Gens in der jeweiligen Wirtszelle ermöglicht, ist dabei bevorzugt. Per Definition wird jede Sequenz oder jedes Gen, das in eine Wirtszelle eingebracht wird, in Bezug auf die Wirtszelle als „heterologe Sequenz“ oder „heterologes Gen“ bezeichnet. Selbst dann, wenn die einzubringende Sequenz oder das einzubringende Gen identisch zu einer endogenen Sequenz oder einem endogenen Gen der Wirtszelle ist. Beispielsweise ist ein Hamster-Aktingen, das in eine Hamster-Wirtszelle eingebracht wird, per Definition ein heterologes Gen.

Bei der rekombinanten Herstellung heterodimerer Proteine, wie z.B. monoklonaler Antikörper (mAk), kann die Transfektion geeigneter Wirtszellen prinzipiell auf zwei verschiedenen Wegen erfolgen. Derartige mAk sind aus mehreren Untereinheiten, den schweren und leichten Ketten, aufgebaut. Für diese Untereinheiten kodierende Gene können in unabhängigen oder in multicistronischen Transkriptionseinheiten auf

einem einzigen Plasmid untergebracht werden, mit dem dann die Wirtszelle transfiziert wird. Dies soll die stöchiometrische Repräsentanz der Gene nach Integration in das Genom der Wirtszelle sichern. Allerdings muss hierbei im Falle unabhängiger Transkriptionseinheiten sichergestellt werden, dass die mRNAs, die für die verschiedenen Proteine kodieren, die gleiche Stabilität, Transkriptions- und Translationseffizienz aufweisen. Im zweiten Fall erfolgt die Expression der Gene innerhalb einer multicistronischen Transkriptionseinheit durch einen einzigen Promotor und es entsteht nur ein Transkript. Durch Verwendung von IRES-Elementen wird eine recht effiziente interne Translationsinitiation der Gene in dem zweiten und den nachfolgenden Cistrons ermöglicht. Dennoch sind die Expressionsraten für diese Cistrons geringer als die des ersten Cistrons, dessen Translationsinitiation über einen sogenannten „cap“-abhängigen Prä-Initiationskomplex wesentlich effizienter ist als die IRES-abhängige Translationsinitiation. Um eine tatsächlich äquimolare Expression der Cistrons zu erreichen, können beispielsweise noch zusätzliche intercistronische Elemente eingeführt werden, die im Zusammenwirken mit den IRES-Elementen für einheitliche Expressionsraten sorgen (WO 94/05785).

Eine andere und erfindungsgemäß bevorzugte Möglichkeit der simultanen Herstellung mehrerer heterologer Proteine ist die Co-Transfektion, bei der die Gene getrennt in verschiedene Expressionsvektoren integriert werden. Dies hat den Vorteil, dass bestimmte Verhältnisse der Gene und Genprodukte zueinander eingestellt werden können, wodurch Unterschiede in der mRNA-Stabilität sowie in der Transkriptions- und Translationseffizienz ausgeglichen werden können. Außerdem sind die Expressionsvektoren wegen ihrer geringeren Größe stabiler und sowohl bei der Klonierung als auch bei der Transfektion einfacher handhabbar.

In einer besonderen Ausführungsform der Erfindung werden daher die Wirtszellen zusätzlich mit einem oder mehreren Vektoren mit Genen, die für ein oder mehrere andere Proteine von Interesse kodieren, transfiziert, bevorzugt co-transfiziert. Der oder die zur Co-Transfektion verwendeten weiteren Vektoren kodieren z.B. für das oder die anderen Proteine von Interesse unter der Kontrolle der gleichen

Promotor/Enhancer-Kombination sowie für zumindest einen weiteren Selektionsmarker, beispielsweise Neomycin-Phosphotransferase.

Erfindungsgemäß werden die Wirtszellen vorzugsweise unter serumfreien 5 Bedingungen etabliert, adaptiert und kultiviert, gegebenenfalls in Medien, die frei von tierischen Proteinen/Peptiden sind. Beispiele für im Handel erhältliche Medien sind Ham's F12 (Sigma, Deisenhofen, DE), RPMI-1640 (Sigma), Dulbecco's Modified Eagle's Medium (DMEM; Sigma), Minimal Essential Medium (MEM; Sigma), Iscove's Modified Dulbecco's Medium (IMDM; Sigma), CD-CHO (Invitrogen, Carlsbad, Ca., 10 USA), CHO-S-SFMII (Invitrogen), serumfreies CHO-Medium (Sigma) und proteinfreies CHO-Medium (Sigma). Jedes dieser Medien kann gegebenenfalls mit verschiedenen Verbindungen ergänzt werden, z.B. Hormonen und/oder anderen Wachstumsfaktoren (z.B. Insulin, Transferrin, epidermalem Wachstumsfaktor, 15 Insulin-ähnlichem Wachstumsfaktor), Salzen (z.B. Natriumchlorid, Calcium, Magnesium, Phosphat), Puffern (z.B. HEPES), Nukleosiden (z.B. Adenosin, Thymidin), Glutamin, Glukose oder anderen äquivalenten Nährstoffen, Antibiotika und/oder Spurenelementen. Obwohl erfindungsgemäß serumfreie Medien bevorzugt sind, können zur Züchtung der Wirtszellen auch Medien verwendet werden, die mit einer geeigneten Menge an Serum versetzt wurden. Zur Selektion von genetisch 20 modifizierten Zellen, die ein oder mehrere Selektionsmarkergene exprimieren, wird dem Medium ein oder mehrere geeignete Selektionsmittel zugefügt.

Als „Selektionsmittel“ wird eine Substanz bezeichnet, die das Wachstum oder das Überleben von Wirtszellen mit einer Defizienz für das jeweilige Selektionsmarkergen 25 beeinträchtigt. Beispielsweise wird zur Selektion auf die Anwesenheit eines exprimierten Antibiotikaresistenzgens wie z.B. der Neomycin-Phosphotransferase das Antibiotikum Geneticin (G418) als Mediumzusatz verwendet. Das Selektionsmittel kann auch ein Stoff sein, der eine Amplifikation des Selektionsmarkergens auslöst, wenn es sich beim dem verwendeten Gen um einen 30 amplifizierbaren Selektionsmarker handelt (siehe Tabelle 1). Methotrexat ist z.B. ein Selektionsmittel, das sich zur Amplifikation des DHFR-Gens eignet. Beispiele für andere amplifikationsauslösende Selektionsmittel sind in der Tabelle 1 aufgeführt.

Ein „Selektionsmarkergen“ ist ein Gen, das die spezifische Selektion von Zellen, die dieses Gen erhalten, durch Zugabe eines entsprechenden Selektionsmittels in das Kultivierungsmedium ermöglicht. Zur Verdeutlichung, ein Antibiotika-Resistenzgen kann als positiver Selektionsmarker verwendet werden. Nur Zellen, die mit diesem 5 Gen transformiert wurden, können in der Gegenwart des entsprechenden Antibiotikums wachsen und somit selektiert werden. Nichtrtransfomierte Zellen hingegen können unter diesen Selektionsbedingungen nicht wachsen oder überleben. Es gibt positive, negative und bifunktionale Selektionsmarker. Positive Selektionsmarker ermöglichen die Selektion und damit Anreicherung von 10 transformierten Zellen durch Vermittlung einer Resistenz gegenüber dem Selektionsmittel oder durch Kompensierung eines metabolischen oder katabolischen Defekts der Wirtszelle. Im Gegensatz dazu können durch negative Selektionsmarker Zellen, die das Gen für den Selektionsmarker erhalten haben, selektiv eliminiert werden. Ein Beispiel hierfür ist das Thymidinkinasegen des Herpes simplex Virus, 15 dessen Expression in Zellen bei gleichzeitiger Gabe von Acyclovir oder Gancyclovir zu deren Elimination führt. Die in dieser Erfindung verwendeten Selektionsmarker, einschließlich der amplifzierbaren Selektionsmarker, schließt gentechnologisch veränderte Mutanten und Varianten, Fragmente, funktionelle Äquivalente, Derivate, Homologe und Fusionen mit anderen Proteinen oder Peptiden ein, solange der 20 Selektionsmarker seine selektiven Eigenschaften beibehält. Solche Derivate weisen eine beträchtliche Homologie in der Aminosäuresequenz in den Bereichen oder Domänen auf, denen die selektive Eigenschaft zugeschrieben wird. In der Literatur 25 sind eine Vielzahl von Selektionsmarkergenen, einschließlich bifunktionaler (positiv/negativ) Marker, beschrieben (siehe z.B. WO 92/08796 und WO 94/28143). Beispiele für Selektionsmarker, die für gewöhnlich in eukaryontischen Zellen verwendet werden, beinhalten die Gene für Aminoglykosid-Phosphotransferase (APH), Hygromycin-Phosphotransferase (HYG), Dihydrofolat-Reduktase (DHFR), Thymidinkinase (TK), Glutamin-Synthetase, Asparagin-Synthetase und Gene, die Resistenz gegenüber Neomycin (G418), Puromycin, Histidinol D, Bleomycin, 30 Phleomycin und Zeocin vermitteln.

Eine Selektion von transformierten Zellen ist auch durch fluoreszenzaktivierte Zellsortierung (FACS) möglich. Hierzu werden beispielsweise bakterielle β -

Galaktosidase, Zelloberflächenmarker oder fluoreszierende Proteine (z.B. grünes fluoreszierendes Protein (GFP) und dessen Varianten von *Aequorea victoria* und *Renilla reniformis* oder anderer Spezies, rote fluoreszierende Protein und in anderen Farben fluoreszierende Proteine und deren Varianten von nicht-biolumineszierenden

5 Organismen wie z.B. *Discosoma sp.*, *Anemonia sp.*, *Clavularia sp.*, *Zoanthus sp.*) zur Selektion von transformierten Zellen eingesetzt.

In der vorliegenden Erfindung wird für die Selektion von genetisch modifizierten (rekombinanten) Wirtszellen als amplifizierbares Selektionsmarker gen der Einsatz 10 des DHFR-Gens bevorzugt. Dieser Marker ist bei der Verwendung von DHFR-negativen Grundzellen wie CHO-DG44 oder CHO-DUKX besonders gut zur Selektion und nachfolgenden Amplifikation geeignet, da diese Zellen kein endogenes DHFR exprimieren und somit nicht in Purin-freiem Medium wachsen. Deshalb kann hier das 15 DHFR-Gen als dominanter Selektionsmarker eingesetzt werden und die transformierten Zellen werden in Hypoxanthin/Thymidin-freiem Medium selektiert. Zur Erzielung einer DHFR-vermittelten Genamplifikation wird Methotrexat (MTX) eingesetzt. Die Wachstumseigenschaften werden maßgeblich durch die Zugabe von MTX beeinflusst. Hierbei ist üblicherweise eine wesentliche Verschlechterung der 20 Fermentationsrobustheit der Zellen mit zunehmender MTX-Konzentration und Amplifikationsstufe zu beobachten. Überraschenderweise wurde jedoch gefunden, dass über das erfindungsgemäße Klonselektionssystem rekombinante Wirtszellen 25 angereichert werden können, die ein erheblich robusteres Verhalten gegenüber hohen MTX-Konzentrationen zeigen (siehe Abb. 7). So konnten Wirtszellen, die mit Hilfe eines Fluoreszenz-aktivierten Zellsortiergeräts (FACS) identifiziert und aussortiert wurden, in Gegenwart von 500 nM, vorzugsweise in Gegenwart von 1 μ M 30 MTX, kultiviert und amplifiziert werden, was zu einer deutlichen Steigerung der Produktivität führte. Somit gilt ein Verfahren zur Selektion von hochproduzierenden Wirtszellen als besonders erfindungsgemäß, bei dem Wirtszellen, die mit einem erfindungsgemäßen Expressionsvektor transfiziert werden und zumindest das Gen von Interesse, das fluoreszierende Protein und ein DHFR-Gen exprimieren, über FACS-Sortierung aussortiert werden, und zumindest einem Genamplifikationsschritt 35 in Gegenwart von zumindest 500 nM, vorzugsweise 1 μ M MTX unterzogen werden.

Unter „Fermentationsrobustheit“ versteht man hierbei Wachstumseigenschaften der Zellen wie z.B. das Einhalten bestimmter Wachstumsraten, Robustheit gegenüber „UpScaling“ (größere Dimensionierung der Bioreaktoren) und Erreichen hoher Zellzahlen und Vitalitäten in der Stammhaltung, um den industriellen

5 Passagiergraten beim „UpScaling“ gerecht zu werden.

Expression

Der Ausdruck „Expression“ bezieht sich auf die Transkription und/oder Translation einer heterologen Gensequenz in einer Wirtszelle. Die Expressionsrate kann hierbei allgemein bestimmt werden, entweder auf Basis der Menge der entsprechenden mRNA, die in der Wirtszelle vorhanden ist, oder auf Basis der produzierten Menge an Genprodukt, das von dem Gen von Interesse kodiert wird. Die Menge der durch Transkription einer ausgewählten Nukleotidsequenz erzeugten mRNA kann 15 beispielsweise durch Northern Blot Hybridisierung, Ribonuklease-RNA-Protektion, in situ Hybridisierung von zellulärer RNA oder durch PCR-Methoden bestimmt werden (Sambrook et al., 1989; Ausubel et al., 1994). Proteine, die von einer ausgewählten Nukleotidsequenz kodiert werden, können ebenfalls durch verschiedene Methoden, wie z.B. durch ELISA, Western Blot, Radioimmunassay, Immunpräzipitation, 20 Nachweis der biologischen Aktivität des Proteins oder durch Immunfärbung des Proteins mit nachfolgender FACS-Analyse, bestimmt werden (Sambrook et al., 1989; Ausubel et al., 1994).

Mit „hohem Expressionslevel (-rate)“, „hoher Expression“, „verstärkter Expression“ 25 oder „hoher Produktivität“ wird die anhaltende und ausreichend hohe Expression oder Synthese einer in eine Wirtszelle eingebrachten heterologen Sequenz, beispielsweise eines für ein therapeutisches Protein kodierenden Gens, bezeichnet. Eine verstärkte oder hohe Expression bzw. ein(e) hohe(r) Expressionslevel (-rate) oder eine hohe Produktivität liegt vor, wenn eine erfindungsgemäße Zelle nach 30 einem hier beschriebenen erfindungsgemäßen Verfahren kultiviert wird, und wenn diese Zelle zumindest mehr als ungefähr 5 pg des gewünschten Genprodukts pro Tag produziert (5 pg/Tag/Zelle). Eine verstärkte oder hohe Expression bzw. ein(e) hohe(r) Expressionslevel (-rate) oder eine hohe Produktivität liegt auch dann vor,

wenn die erfindungsgemäße Zelle zumindest mehr als ungefähr 10 pg des gewünschten Genprodukts pro Tag produziert (10 pg/Tag/Zelle). Eine verstärkte oder hohe Expression bzw. ein(e) hohe(r) Expressionslevel (-rate) oder eine hohe Produktivität liegt insbesondere auch vor, wenn die erfindungsgemäße Zelle 5 zumindest mehr als ungefähr 15 pg des gewünschten Genprodukts pro Tag produziert (15 pg/Tag/Zelle). Eine verstärkte oder hohe Expression bzw. ein(e) hohe(r) Expressionslevel (-rate) oder eine hohe Produktivität liegt insbesondere dann vor, wenn die erfindungsgemäße Zelle zumindest mehr als ungefähr 20 pg des gewünschten Genprodukts pro Tag produziert (20 pg/Tag/Zelle). Eine besondes 10 verstärkte oder hohe Expression bzw. ein(e) besonders hohe(r) Expressionslevel (-rate) oder eine besonders hohe Produktivität liegt dann vor, wenn die erfindungsgemäße Zelle zumindest mehr als ungefähr 30 pg des gewünschten Genprodukts pro Tag produziert (30 pg/Tag/Zelle).

15 Eine hohe oder verstärkte Expression, eine hohe Produktivität oder ein(e) hohe(r) Expressionslevel (-rate) im Sinne der vorliegenden Erfindung kann auf verschiedene Weisen erzielt werden. Beispielsweise können durch die Co-Expression des Gens von Interesse mit einem Gen für einen amplifizierbaren Selektionsmarker Zellen selektioniert und identifiziert werden, die das heterologe Gen in hohem Maße 20 exprimieren. Der amplifizierbare Selektionsmarker ermöglicht dabei nicht nur die Selektion von stabil transfizierten Wirtszellen sondern auch die Genamplifikation des heterologen Gens von Interesse. Die Integration der zusätzlichen Kopien der 25 Nukleinsäuren kann dabei in das Genom der Wirtszellen, in zusätzliche artifizielle/Mini-Chromosomen oder in episomal lokalisierte Polynukleotide erfolgen. Diese Vorgehensweise kann mit einer durch FACS gestützten Selektion von 30 rekombinanten Wirtszellen, die als weiteren Selektionsmarker beispielsweise ein (oder mehrere) Fluoreszenzprotein(e) (z.B. GFP) oder einen Zelloberflächenmarker enthalten, kombiniert werden. Andere Methoden zur Erzielung einer verstärkten Expression, wobei auch eine Kombination verschiedener Methoden möglich ist, beruhen beispielsweise auf der Verwendung von (artifiziellen) Transkriptionsfaktoren, Behandlung der Zellen mit natürlichen oder synthetischen Agentien zur Hochregulation endogener oder heterologer Genexpression, Verbesserung der Stabilität (Halbwertszeit) der mRNA oder des Proteins, Verbesserung der mRNA-

Translationsinitiation, Erhöhung der Gendosis durch Verwendung von episomalen Plasmiden (basierend auf der Verwendung von viralen Sequenzen als Replikationsursprung, z.B. von SV40, Polyoma, Adenovirus, EBV oder BPV), Einsatz von amplifikationsfördernden Sequenzen (Hemann et al., 1994) oder auf DNA-Konkaternen basierenden *in vitro* Amplifikationssystemen (Monaco et al., 1996).

Erfindungsgemäß ist eine gekoppelte Transkription des Gens von Interesse und des Gens, das für das fluoreszierende Protein kodiert. Von der resultierenden bicistronischen mRNA werden sowohl das Protein von Interesse als auch das fluoreszierende Protein exprimiert. Aufgrund dieser Kopplung der Expression des Proteins von Interesse und des fluoreszierenden Proteins ist es erfindungsgemäß leicht möglich, hochproduzierende rekombinante Wirtszellen über das exprimierte fluoreszierende Protein zu selektionieren und zu isolieren, z.B. durch Sortieren mit Hilfe eines Fluoreszenz-aktivierten Zellsortiergeräts (FACS).

Die Selektion von rekombinanten Wirtszellen, die eine hohe Vitalität und eine erhöhte Expressionsrate des gewünschten Genprodukts zeigen, ist ein mehrstufiger Prozess. Die mit dem erfindungsgemäßen Expressionsvektor transfizierten oder gegebenenfalls einem weiteren Vektor z.B. co-transfizierten Wirtszellen werden zumindest auf die Expression des mit dem Gen von Interesse gekoppelten Gens, das für ein fluoreszierendes Protein kodiert, untersucht, um die Zellen/Zellpopulation zu identifizieren und zu selektionieren, die die höchsten Expressionsraten an fluoreszierendem Protein zeigen. Vorzugsweise werden nur die Zellen aussortiert und weiter kultiviert, die zu den 10% Zellen mit der höchsten Expressionsrate an fluoreszierendem Protein gehören. In der Praxis bedeutet dies, dass die hellsten 10% der fluoreszierenden Zellen heraussortiert und weiter kultiviert werden. Entsprechend können auch die hellsten 5%, vorzugsweise die hellsten 3%, oder auch nur die hellsten 1% der fluoreszierenden Zellen eines Zellgemisches heraussortiert und vermehrt werden. In einer besonders bevorzugten Ausführungsform werden lediglich die hellsten 0.5% bzw. die hellsten 0,1% der fluoreszierenden Zellen heraussortiert und vermehrt.

Hierzu kultiviert man die zuvor mit dem erfindungsgemäßen Expressionsvektor transformierten Zellen in einem Selektionsmedium, das gegebenenfalls auch ein für den amplifizierbaren Selektionsmarker spezifisches Selektionsmittel enthält. Dabei können schrittweise erhöhte Konzentrationen an Selektionsmittel verwendet werden, 5 um einen stufenweise erhöhten Selektionsdruck auszuüben.

Der Selektionsschritt kann an Zellpools oder mit bereits vorsortierten Zellpools/ Zellklonen durchgeführt werden. Es können ein oder mehrere, vorzugsweise zwei oder mehr und insbesondere drei oder mehr Sortierungsschritte durchgeführt 10 werden, wobei zwischen den einzelnen Sortierungsschritten die Zellen über einen bestimmten Zeitraum, z.B. etwa zwei Wochen bei Pools, kultiviert und vermehrt werden.

Gegebenenfalls kann man die Wirtszellen einem oder mehreren Genamplifikations- 15 schritten unterziehen, um die Kopienzahl zumindest des Gens von Interesse und des amplifizierbaren selektionierbaren Markgens zu erhöhen. Verfahren zur stufenweise Genamplifikation mit Hilfe von Methotrexat sind beispielhaft in US 5,179,017 beschrieben. Erfindungsgemäß ist die erzielbare hohe Produktivität nicht an eine erhöhte Anzahl von Genkopien gebunden. Sie ist vielmehr Ausdruck einer erhöhten 20 Stabilität und Fermentationsrobustheit der Hochleistungsklone. Es ist daher möglich, die Anzahl der erforderlichen Genamplifikationsschritte zu reduzieren und z.B. nur eine einzige Genamplifikation durchzuführen.

Demnach ist ein Verfahren zur Selektion von Zellen erfindungsgemäß, das folgende 25 Schritte enthält:

- i) Transformation geeigneter Wirtszellen zumindest mit einem der erfindungsgemäßen Vektoren, wobei die DNA der Expressionsvektoren vorzugsweise stabil in das Wirtszellgenom oder in artifizielle Chromosomen/Minichromosomen eingebaut wird;
- 30 ii) die transformierten Zellen unter Bedingungen kultiviert werden, die eine Expression des Gens von Interesse und des fluoreszierenden Proteins erlauben;

iii) die Zellen in Anwesenheit zumindest eines Selektionsmittels kultiviert werden, so dass nur solche Zellen vermehrt werden, die in Gegenwart des besagten Selektionsmittel wachsen können;

iv) das Aussortieren von Zellen aus einem Zellgemisch, die die höchste Expressionsrate an fluoreszierendem Protein zeigen, wobei die Zellen mit Hilfe eines Fluoreszenz-aktivierten Zellsortiergeräts (FACS) detektiert und sortiert werden;

v) die Kultivierung der aussortierten Zellen mit den höchsten Expressionsraten für das fluoreszierende Protein.

10 Optional können die Schritte ii) – v) mit den nach Schritt v) gewonnenen Zellen einfach oder mehrfach wiederholt werden. Ferner können die transformierten Zellen optional auch zusätzlich einem oder mehreren Genamplifikationsschritten unterzogen werden, in dem sie in Gegenwart eines Selektionsmittels kultiviert werden, dass zu einer Amplifikation des amplifizierbaren selektierbaren Markergens führt. Dieser 15 Schritt kann sowohl mit noch nicht sortierten Zellen als auch mit bereits einfach oder mehrfach vorsortierten Zellen erfolgen.

Weiterhin erfindungsgemäß ist ein Verfahren, bei dem entsprechend aussortierte Zellen vermehrt werden und zur Herstellung des kodierenden Genprodukt von 20 Interesse verwendet wird. Hierzu werden die selektionierten hoch-produzierenden Zellen vorzugsweise in einem serumfreien Kulturmedium und vorzugsweise in Suspensionskultur unter Bedingungen gezüchtet, die eine Expression des Gens von Interesse erlauben. Das Protein von Interesse wird dabei vorzugsweise als sekretiertes Genprodukt aus dem Zellkulturmedium gewonnen. Bei Expression des 25 Proteins ohne Sekretionssignal kann das Genprodukt aber auch aus Zelllysaten isoliert werden. Um ein reines, homogenes Produkt zu erhalten, das im Wesentlichen frei ist von anderen rekombinanten Proteinen und Wirtszellproteinen, werden übliche Reinigungsschritte durchgeführt. Hierzu entfernt man häufig zunächst Zellen und Zelltrümmer aus dem Kulturmedium oder Lysat. Das gewünschte Genprodukt kann 30 dann von kontaminierenden löslichen Proteinen, Polypeptiden und Nukleinsäuren befreit werden, z.B. durch Fraktionierung an Immunaffinitäts- und Ionenaustauschsäulen, Ethanolfällung, Umkehrphasen-HPLC oder Chromatographie an Sephadex, Silica oder Kationenaustauscherharzen wie DEAE. Methoden, die zur

Aufreinigung eines von rekombinanten Wirtszellen exprimierten heterologen Proteins führen, sind dem Fachmann bekannt und sind in der Literatur beschrieben, z.B. bei Harris et al. (1995) und Scopes (1988).

5 Im folgenden wird die Erfindung anhand nicht-beschränkender Ausführungsbeispiele näher erläutert.

BEISPIELE

10

Abkürzungen

AP:	alkalische Phosphatase
bp:	Basenpaar
15 CHO:	Chinese hamster ovary
DHFR:	Dihydrofolat-Reduktase
ELISA:	enzyme-linked immunosorbant assay
FACS:	fluorescence-activated cell sorter
20 FAP:	Fibroblasten-aktiviertes Protein
GFP:	grünes fluoreszierendes Protein
HBSS:	Hanks Balanced Salt Solution
HT:	Hypoxanthin/Thymidin
HRPO:	horseradish peroxidase
25 IRES:	internal ribosomal entry site
kb:	Kilobase
mAk:	monoklonaler Antikörper
MTX:	Methotrexat
PCR:	polymerase chain reaction
30 sICAM:	soluble intracellular adhesion molecule

30

Methoden

1. Zellkultur und Transfektion

Die Zellen CHO-DG44/DHFR-/ (Urlaub et al., 1983) wurden permanent als 35 Suspensionzellen in serum-freiem und mit Hypoxanthin und Thymidin supplementiertem CHO-S-SFMII Medium (Invitrogen GmbH, Karlsruhe, DE) in Zellkulturflaschen bei 37°C in feuchter Atmosphäre und 5% CO₂ kultiviert. Die Zellzahlen sowie die Viabilität wurden mit einem CASY1 Cell Counter (Schaerfe

System, DE) oder durch Trypanblau-Färbung bestimmt und die Zellen dann in einer Konzentration von $1 - 3 \times 10^5$ /mL eingesät und alle 2 – 3 Tage passagiert.

Zur Transfektion von CHO-DG44 wurde Lipofectamine Plus Reagenz (Invitrogen 5 GmbH) eingesetzt. Pro Transfektionsansatz wurden dabei insgesamt 1 μ g Plasmid-DNA, 4 μ L Lipofectamine und 6 μ L Plus-Reagenz nach den Angaben des Herstellers gemischt und in einem Volumen von 200 μ L zu 6×10^5 exponentiell wachsenden CHO-DG44 Zellen in 0,8 mL HT-supplementiertem CHO-S-SFMII Medium gegeben. Nach dreistündiger Inkubation bei 37°C in einem Zellinkubator erfolgte eine Zugabe 10 von 2 mL HT-supplementiertem CHO-S-SFMII Medium. Zur DHFR-basierten Selektion von stabil transfizierten CHO-DG44 wurden die Zellen 2 Tage nach 15 Transfektion in CHO-S-SFMII Medium ohne Hypoxanthin- und Thymidinzusatz transferiert, wobei das Medium alle 3 bis 4 Tage gewechselt wurde. Bei einer DHFR- und Neomycin-Phosphotransferase-basierten Selektion im Falle einer Co-Transfektion, in der der eine Expressionsvektor einen DHFR- und der andere Expressionsvektor einen Neomycin-Phosphotransferase-Selektionsmarker enthielt, wurde dem Medium außerdem noch G418 (Invitrogen) in einer Konzentration von 400 μ g/mL zugesetzt.

20 Eine DHFR-basierte Genamplifikation der integrierten heterologen Gene wurde durch Zugabe des Selektionsmittels MTX (Sigma, Deisenhofen, DE) in einer Konzentration von 5 – 2000 nM zum HT-freien CHO-S-SFMII Medium erreicht.

2. Expressionsvektoren

25 Zur Expressionsanalyse wurden eukaryontische Expressionsvektoren eingesetzt, die auf dem pAD-CMV Vektor (Werner et al., 1998) basieren und die konstitutive Expression eines heterologen Gens über die Kombination CMV Enhancer/Hamster Ubiquitin/S27a Promotor (WO 97/15664) vermitteln. Während der Basisvektor pBID das DHFR-Minigen enthält, das als amplifizierbarer Selektionsmarker dient (siehe 30 z.B. EP 0 393 438), ist im Vektor pBIN das DHFR-Minigen durch ein Neomycin-resistenzgen ersetzt worden (Abb.2). Hierzu wurde der Selektionsmarker Neomycin-Phosphotransferase, inklusive SV40 early Promotor und TK-Polyadenylierungs-signal, aus dem kommerziellen Plasmid pBK-CMV (Stratagene, La Jolla, CA, USA)

als 1640 bp Bsu36I-Fragment isoliert. Nach einer Auffüllreaktion der Fragmentenden durch Klenow-DNA-Polymerase wurde das Fragment mit dem 3750 bp Bsu36I/StuI-Fragment des Vektors pBID, das ebenfalls mit Klenow-DNA-Polymerase behandelt wurde, ligiert.

5 Im bicistronischen Basisvektor pBIDG (Abb.2) wurde die IRES-GFP-Genregion aus dem Vektor pIRES2-EGFP (Clontech, Palo Alto, CA, USA) isoliert und in einer Weise unter die Kontrolle des CMV Enhancer/Promotors im Vektor pBID gebracht, dass die multiple Klonierungsstelle zwischen Promotorregion und IRES-Element erhalten blieb. Dabei wurde wie folgt vorgegangen. In einer PCR-Mutagenese, wobei das

10 Plasmid pIRES2-EGFP als Template diente, wurde zum einen die HindIII-Schnittstelle AAGCTT innerhalb der IRES-Sequenz durch den Einsatz mutagener
Primer in die Sequenzabfolge ATGCTT umgewandelt und somit entfernt. Zum anderen wurde mittels eines Primers mit Komplementarität zum 5'Ende der IRES-Sequenz eine XbaI-Schnittstelle und bzw. mit Komplementarität zum 3'Ende der

15 GFP-Sequenz eine SphI-Schnittstelle eingeführt. Das resultierende PCR-Fragment, das die komplette IRES- und GFP-Sequenz umfasste, wurde mit XbaI und SphI verdaut und in die singuläre XbaI-Schnittstelle am 3'Ende der multiplen Klonierungsstelle des Vektors pBID kloniert.

20 Das humane sICAM-Gen wurde als HindIII/Sall Fragment aus pAD-sICAM (Werner et al., 1998) isoliert und in die entsprechenden Schnittstellen des Vektors pBIDG kloniert, wobei der Vektor pBIDG-sICAM resultierte (Abb.3).

Zur Expression des monoklonalen humanisierten F19 Antikörpers wurde die schwere Kette als 1,5 kb Nael/HindIII-Fragment aus dem Plasmid pG1D105F19HC (NAGENESEQ: AAZ32786) isoliert und in den mit EcoRI (mit Klenow-DNA-Polymerase aufgefüllt) und HindIII-verdauten Vektor pBIDG kloniert, wobei der Vektor pBIDG-F19HC resultierte (Abb.3). Die leichte Kette hingegen wurde als 1,3 kb HindIII/EcoRI-Fragment aus dem Plasmid pKN100F19LC (NAGENESEQ: AAZ32784) isoliert und in die entsprechenden Schnittstellen des Vektors pBIN kloniert, wodurch der Vektor pBIN-F19LC entstand (Abb.3).

3. FACS

Die flowcytometrischen Analysen und Sorting wurden mit einem Coulter Epics Altra-Gerät durchgeführt. Das FACS ist mit einem Helium-Argon-Laser mit einer Anregungswellenlänge von 488 nm ausgestattet. Die Fluoreszenzintensität wird bei einer dem Fluoreszenzprotein adäquaten Wellenlänge aufgenommen und mittels der angeschlossenen Software Coulter Expo32 prozessiert. Das Sorting wurde mit einer Rate von 8000 – 10000 Events/Sekunde durchgeführt. Die suspendierten Zellen wurden abzentrifugiert (5 min bei 180xg) und in HBSS auf eine Zellkonzentration von 1 – 1,5 x10⁷/mL eingestellt. Anschließend erfolgte ein Sorting der Zellen entsprechend ihres Fluoreszenzprotein-Signals. Die Zellen wurden in Röhrchen mit vorgelegtem Kultivierungsmedium aufgenommen, abzentrifugiert und abhängig von der aussortierten Zellzahl in entsprechende Kultivierungsgefäß eingesät.

4. ELISA

Die sICAM-Titer in Überständen von stabil transfizierten CHO-DG44 Zellen wurden mittels ELISA nach Standardprotokollen quantifiziert (Ausubel et al., 1994, updated), wobei zwei im Haus entwickelte sICAM-spezifische mAk verwendet wurden (z.B. in US 5,284,931 und US 5,475,091 beschrieben). Bei einem der beiden Antikörper handelt es sich dabei um einen HRPO-konjugierten Antikörper. Als Standard wurde gereinigtes sICAM-Protein eingesetzt.

Die Quantifizierung des F19 mAk in den Überständen von stabil transfizierten CHO-DG44 Zellen erfolgte mittels ELISA nach Standardprotokollen (Ausubel et al., 1994, updated), wobei zum einen ein Ziege anti-Human IgG Fc-Fragment (Dianova, Hamburg, DE) und zum anderen ein AP-konjugierter Ziege anti Human Kappa light chain Antikörper (Sigma) eingesetzt wurde. Als Standard diente gereinigter F19 Antikörper. Produktivitäten (pg/Zelle/Tag) wurden dabei mit der Formel pg/((Ct-Co) t / ln (Ct-Co)) berechnet, wobei Co und Ct die Zellzahl bei Aussaat bzw. Ernte und t die Kultivierungsdauer angibt.

Beispiel 1: Vergleich der CMV- und der Hamster-Ubiquitin/S27a-Promotoraktivität

Um die Aktivität des Hamster-Ubiquitin/S27a-Promotors mit der des häufig in eukaryontischen Expressionsvektoren verwendeten CMV-Promotors zu vergleichen, 5 wurden CHO-DG44 Zellen mit diversen rekombinanten Vektoren transfiziert. Das heterologe Genprodukt, zum einen ein lysosomales Enzym, zum anderen ein IgG1-Antikörper, wurde dabei entweder unter der Kontrolle des CMV-Promotors oder unter der des Hamster-Ubiquitin/S27a-Promotors exprimiert. Beide Promotoren waren dabei funktionell mit dem CMV-Enhancer verknüpft. Als Terminationssignal für das 10 heterologe Gen wurde das BGH polyA verwendet. Die Expressionsvektoren, die den CMV-Promotor beinhalteten, basierten dabei entweder auf einem modifizierten pcDNA3- („CMV¹“, Invitrogen) oder pBluescript-Vektor („CMV²“, Stratagene) und kodierten zusätzlich für den amplifizierbaren Selektionsmarker Dihydrofolat-Reduktase. Der Expressionsvektor mit dem Hamster-Promotor basiert hingegen auf 15 dem pAD-CMV Vektor (Werner et al., 1998). Zur Expression der schweren und leichten Kette des Antikörpers wurde eine Co-Transfektion mit einem zweiten Vektor durchgeführt, der als Selektionsmarker ein Neomycin-Resistenzgen enthielt. Der CMV-Enhancer kann aber auch durch den SV40-Enhancer ersetzt werden.

20 Durch limitierte Verdünnung in 96Well-Platten wurden nach der Transfektion Zellklone in HT-freiem Medium (im Falle der Co-Transfektion noch Zusatz von 400 µg/mL G418) selektiert und isoliert. Zellklone mit der höchsten Produktivität hinsichtlich des rekombinanten Proteins wurden einer stufenweisen DHFR-basierten Genamplifikation durch schrittweise Erhöhung der Methotrexat-Konzentration von 5 25 nM über 50 nM, 500 nM auf 2 µM unterzogen, jeweils verbunden mit einer Verdünnungsklonierung. Auf jeder Amplifikationsstufe wurden dabei ca. 20 bis 30 Klone mit der jeweils höchsten Produktivität ausgewählt.

Generell erwies sich der Hamster-Promotor als leistungsstärker. Sowohl bei der 30 Expression des lysosomalen Enzyms als auch bei der Expression des Antikörpers wurden Produktivitäten bzw. Titer erzielt, die um den Faktor 2 bis 5 höher lagen als bei den Zellen, in denen das heterologe Gen unter der Kontrolle des CMV-Promotors exprimiert wurde. In Abbildung 1 sind beispielhaft die relativen Titer und relativen

spezifischen Produktivitäten der jeweils besten Zellklone auf der jeweiligen Amplifikationsstufe dargestellt, wobei die auf den CMV-Promotor basierte Expression für das jeweilige heterologe Gen als 1 gesetzt wurde (CMV¹ für das lysosomale Enzym, CMV² für den Antikörper).

5

Beispiel 2: Isolierung von hochexprimierenden sICAM-Zellen durch GFP-basiertes FACS-Sorting

Die lösliche Form des interzellulären Adhäsionsmoleküls ICAM1, sICAM, ist ein 10 mögliches Therapeutikum bei Erkältungen, da es mit dem ICAM-Rezeptor um die Bindung von Rhinoviren kompetiert und auf diese Weise deren Interaktion mit dem ICAM-Rezeptor, Voraussetzung für den Eintritt in die Zellen und die nachfolgende Infektion (Bella et al., 1999; Marlin et al., 1990), reduzieren oder sogar verhindern kann.

15 sICAM wurde als Beispiel für die Expression eines einzelkettigen Proteins (480 Aminosäuren) in CHO-Zellen gewählt. Dazu wurden CHO-DG44 mit pBIDG-sICAM transfiziert (Abb.3). Die zusätzliche Expression von GFP in pBIDG-sICAM transfizierten Zellen ermöglichte die Anwendung einer FACS-basierten Selektionsstrategie. Das therapeutische Protein sICAM und GFP wurden dabei 20 gemeinsam von einer bicistronischen und das DHFR von einer separaten Transkriptionseinheit exprimiert. Zwei bis drei Wochen nach der ersten Selektion in HT-freiem CHO-S-SFMII Medium wurden die 5% der Zellen mit der höchsten GFP-Fluoreszenz heraussortiert. Nach ca. zweiwöchiger Kultivierung wurden wiederum 25 die 5% Zellen mit der höchsten GFP-Fluoreszenz isoliert. Insgesamt wurde diese sequentielle Sortierung sechsmal durchgeführt. Dabei konnte eine gute Korrelation zwischen sICAM-Produktivität und GFP-Fluoreszenz gezeigt werden (Abb. 4). Allein durch die FACS-gestützte Selektion, ohne jeglichen MTX-Amplifikationsschritt, wurden so in kürzester Zeit Zellpools mit hohen spezifischen Produktivitäten von bis 30 zu 16 pg/Zelle/Tag isoliert (Abb. 5). Durch Kombination der GFP-basierten Selektion mit einem einzigen nachfolgenden MTX-Amplifikationsschritt war sogar eine Steigerung der Produktivität auf über 30 pg/Zelle/Tag möglich (Abb. 6). Diese Produktivitäten wurden sowohl bei einer Amplifikation eines Pools nach dem vierten Sorting mit 500 nM MTX als auch bei der Amplifikation eines Pools nach dem

sechsten Sorting mit 2 μ M MTX erreicht. Im Gegensatz zu einer stufenweise Amplifikation, bei der in der Regel mit sehr geringen MTX-Konzentrationen im Bereich von 5 – 20 nM MTX begonnen wird, musste hier zur Erzielung eines Amplifikationseffekts eine von Anfang an höhere MTX-Konzentration eingesetzt werden. So resultierte aus der Zugabe von 5 oder 50 nM MTX zu Zellen aus dem vierten Sorting beziehungsweise 500 nM zu Zellen aus dem sechsten Sorting keine signifikante Steigerung der Produktivität (Abb. 6). Offensichtlich war der in den Ausgangspools vorliegende Level an DHFR schon so hoch, dass nur mit einer hohen MTX-Dosis eine komplette DHFR-Inhibierung erzielt werden konnte. Außerdem überstanden die vorsortierten Zellpools trotz der hohen initialen MTX-Dosis die Selektionsphase viel besser, d.h. es wurde in kürzerer Zeit wieder eine Zellpopulation mit hoher Vitalität erhalten als bei der konventionellen stufenweisen Genamplifikationsstrategie (Abb. 7).

15 **Beispiel 3:** Isolierung von Zellen mit hoher Expression des mAks F19 durch GFP-basiertes FACS-Sorting

In einer Co-Transfektion wurden CHO-DG44 Zellen mit der Plasmidkombination pBIDG-F19HC und pBIN-F19LC transfiziert (Abb.3). Der exprimierte humanisierte 20 Antikörper F19 ist dabei gegen das Oberflächenmolekül FAP gerichtet, das von reaktiven Stroma-Fibroblasten synthetisiert wird (siehe dazu auch Patent EP 0 953 639). In den eingesetzten Vektorkonfigurationen werden die beiden Proteinketten des Antikörpers jeweils von einem eigenen Vektor, der zusätzlich noch für einen DHFR- bzw. Neomycin-Phosphotransferase-Selektionsmarker in einer 25 separaten Transkriptionseinheit kodiert, exprimiert. Zusätzlich ist im Vektor pBIDG-F19HC ein weiterer Selektionsmarker, das GFP, enthalten. Durch die transkriptionelle Verknüpfung der Expression des GFPs und der schweren Kette mittels eines IRES-Elements konnten bei der Co-Transfektion von CHO-DG44 mit den Vektoren pBIDG-F19HC/pBIN-F19LC in kurzer Zeit Zellen mit einer hohen 30 Expression des Antikörpers F19 allein dadurch isoliert werden, dass mittels sequentiellem FACS-Sorting die Zellen mit einem hohen GFP-Gehalt selektiert wurden. Hierzu wurden nach einer ersten zwei- bis dreiwöchigen Selektion der transfizierten Zellpools in HT-freiem CHO-S-SFMII Medium mit Zusatz von 400

5 $\mu\text{g/mL}$ G418 die 5% der Zellen mit der höchsten GFP-Fluoreszenz durch FACS heraussortiert. Dieses Sorting wurde insgesamt bis zu sechsmal durchgeführt, wobei zwischen jedem Sorting eine Kultivierungsperiode von ca. 2 Wochen lag. Erstaunlicherweise konnte dabei eine gute Korrelation zwischen F19-Produktivität und GFP-Fluoreszenz gezeigt werden (Abb. 8), obwohl beide Proteinketten jeweils von einem eigenen Vektor aus exprimiert wurden und beim GFP-basierten FACS-Sorting zudem auch nur auf die Expression der schweren Kette, bedingt durch deren transkriptionelle Kopplung mit GFP, selektiert werden konnte. Die Produktivitäten konnten auf bis zu 10 pg/Zelle/Tag gesteigert (Abb. 9) und durch einen einzigen 10 nachfolgenden MTX-Amplifikationsschritt, ausgehend von dem Zellpool aus dem fünften Sorting, durch Zugabe von 1000 nM MTX zum Selektionsmedium noch auf durchschnittlich 37 pg/Zelle/Tag erhöht werden. Vergleichbare Daten konnten auch bei einer funktionellen Verknüpfung des Hamster-Promotors mit dem SV40-Enhancer anstelle des CMV-Enhancers erzielt werden. Gleichzeitig konnte dadurch auch die 15 Entwicklungszeit zur Selektion von hoch-produzierenden Zellen im Vergleich zu einer konventionellen stufenweisen Genamplifikationsstrategie, die in der Regel 4 Stufen umfasst, um die Hälfte auf ca. 120 Tage reduziert werden, einhergehend mit einer signifikanten Reduktion der Entwicklungskapazitäten und –kosten.

20

LITERATUR

25 Adam, M.A. et al., *J Virol* **1991**, 65, 4985 - 4990
Ausubel, F.M. et al., *Current Protocols in molecular biology*. New York : Greene Publishing
Associates and Wiley-Interscience. **1994** (updated)

Bella, J. et al., *J. Struct. Biol.* **1999**, 128, 69 - 74
Bennett, R.P. et al., *BioTechniques* **1998**, 24, 478 – 482
Chalfie, M. et al., *Science* **1994**, 263, 802 - 805
Chamov, S.M. et al., *Antibody Fusion Proteins*, Wiley-Liss Inc., **1999**

30 Davies, M.V. et al., *J Virol* **1992**, 66, 1924 – 1932
Faisst, S. et al., *Nucleic Acids Research* **1992**, 20, 3 – 26
Gossen, M. et al., *Curr Op Biotech* **1994**, 5, 516 - 520
Gubin, A.N. et al., *Biochem Biophys Res Commun* **1997**, 236, 347 – 350

Haber, D.A. et al., *Somatic Cell Genetics* **1982**, 8, 499 - 508

Harris et al., Protein Purification : A Practical Approach, Pickwood and Hames, eds., IRL Press, **1995**

Hemann, C. et al., *DNA Cell Biol* **1994**, 13 (4), 437 - 445

5 Huston, C. et al., *Proc Natl Acad Sci USA* **1988**, 85 (16), 5879 - 5883

Hu, S. et al., *Cancer Res.* **1996**, 56 (13), 3055 - 3061

Jang, S.K. et al., *J Virol* **1989**, 63, 1651 - 1660

Kaufman, R.J. et al., *J Mol Biol* **1982**, 159, 601 - 621

Kaufman, R.J., *Methods in Enzymology* **1990**, 185, 537 - 566

10 Kortt, A.A. et al., *Protein Engineering* **1997**, 10 (4), 423 - 433

Levenson, V.V. et al., *Human Gene Therapy* **1998**, 9, 1233 - 1236

Lottspeich F. and Zorbas H. eds., Bioanalytic, Spektrum Akad. Verl., 1998

Lovejoy, B. et al., *Science* **1993**, 259, 1288 - 1293

Marlin, S.D. et al., *Nature* **1990**, 344, 70 - 77

15 Meng, Y.G. et al., *Gene* **2000**, 242, 201 - 207

Monaco, L. et al., *Gene* **1996**, 180, 145 - 150

Morgan, R.A. et al., *Nucleic Acids Research* **1992**, 20, 1293 - 1299

Mosser, D.D. et al., *BioTechniques* **1997**, 22, 150 - 161

Nolan, G.P. et al., *Proc Natl Acad Sci USA* **1988**, 85, 2603 - 2607

20 Ohshima, Y. et al., *J Mol Biol* **1987**, 195, 247 - 259

Pack, P. et al., *Biotechnology* **1993**, 11, 1271 - 1277

Pack, P. et al., *J Mol Biol* **1995**, 246 (11), 28 - 34

Pelletier, J. et al., *Nature* **1988**, 334, 320 - 325

Perisic, O. et al., *Structure* **1994**, 2, 1217 - 1226

25 Primig, M. et al., *Gene* **1998**, 215, 181 - 189

Ramesh, N. et al., *Nucleic Acids Research* **1996**, 24, 2697 - 2700

Sambrook, J., Fritsch, E.F. & Maniatis, T., Molecular Cloning: A Laboratory Manual Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, **1989**

Scopes, R., Protein Purification, Springer Verlag, **1988**

30 Simonson, C.C. et al., *Proc Natl Acad Sci USA* **1983**, 80, 2495 - 2499

Sugimoto et al., *Biotechnology* **1994**, 12, 694 - 698

Urlaub, G. et al., *Cell* **1983**, 33, 405 - 412

Werner, R.G. et al., *Arzneim.-Forsch./Drug.Res.* **1998**, 48, 870 - 880

Wigler, M. et al., *Proc Natl Acad Sci USA* **1980**, 77, 3567 – 3570

	aggggcagga	gggggctaca	cggaagaggc	cacacccgca	cttgggaaga	ctcgatttg	1740
	gcttcagctg	gctgagacgc	cccagcaggc	tcctcggtca	caccttcagc	cccgaaatgc	1800
	ttccggccca	taacccttcc	cttctaggca	tttccggcga	ggaccccaccc	tcgcgc当地	1860
	cattcggccc	catccccgg	tcctcacctg	aatctctaac	tctgactcca	gagtttagag	1920
5	actataacca	gataggccgg	atgtgtggaa	ctgcacatcttgc	ggacgagtag	tttttagcaaa	1980
	aagaaaagcga	cgaaaaacta	caattcccgag	acagacttgt	gttacaccttc	ttctcatgtct	2040
	aaacaagccc	cctttaaagg	aaagcccctc	tttagtcgcat	cgactgtgt	agaaaaggcgt	2100
	ttgaaacatt	ttaatgttgg	gcacaccgtt	tcgaggaccg	aatgagaaa	gagcataggg	2160
	aaacggagcg	cccgagctag	tctggcaactg	cgtagacag	ccgcggcgt	tgcaaggccgc	2220
10	aggcacttgc	gtggacgcct	aaggggcggg	tctttcggcc	ggaagcccc	gttggccgc	2280
	gcggctcttc	cttccgatc	cgccatccgt	ggtagtgt	tgctgcgggc	tgccgc当地	2340
	gcttggggct	tcccgctcg	ctctcacccct	ggtcggcggc	tctaattccgt	ctctttcga	2400
	atgttag						2406

15

PATENTANSPRÜCHE

1. Expressionsvektor umfassend ein Gen, das für ein Protein von Interesse kodiert, in funktioneller Verknüpfung mit einem Hamster-Ubiquitin/S27a-Promotor und einem Gen, das für ein fluoreszierendes Protein kodiert.
5
2. Expressionsvektor nach Anspruch 1, dadurch gekennzeichnet, dass er ein amplifizierbares Selektionsmarkergen enthält.
- 10 3. Expressionsvektor nach Anspruch 1 oder 2, dadurch gekennzeichnet, dass er einen oder mehrere Enhancer in funktioneller Verknüpfung mit dem oder den Promotoren umfasst.
- 15 4. Expressionsvektor nach einem der Ansprüche 1 bis 3, dadurch gekennzeichnet, dass er zusätzlich eine interne Ribosomenbindungsstelle (IRES) enthält, welche die bicistronische Expression des Gens, das für ein fluoreszierendes Protein kodiert, und des Gens, das für ein Protein/Produkt von Interesse kodiert, ermöglicht.
- 20 5. Expressionsvektor nach einem der Ansprüche 1 bis 4, dadurch gekennzeichnet, dass sich das Gen, das für ein fluoreszierendes Protein kodiert, und das amplifizierbare Selektionsmarkergen in einer oder in zwei separaten Transkriptionseinheiten befinden.
- 25 6. Expressionsvektor nach einem der Ansprüche 1 bis 5, dadurch gekennzeichnet, dass die funktionelle Verknüpfung nicht über Intronsequenzen erfolgt.
- 30 7. Expressionsvektor nach einem der Ansprüche 1 bis 6, dadurch gekennzeichnet, dass das amplifizierbare Selektionsmarkergen für Dihydrofolat-Reduktase (DHFR) oder ein Fusionsprotein aus dem fluoreszierenden Protein und DHFR kodiert.

8. Expressionsvektor nach einem der Ansprüche 3 bis 7, dadurch gekennzeichnet, dass der Enhancer ein CMV- oder SV40-Enhancer ist.
- 5 9. Expressionsvektor nach einem der Ansprüche 1 bis 8, dadurch gekennzeichnet, dass er zusätzlich zumindest ein Polyadenylierungssignal aufweist.
- 10 10. Expressionsvektor nach einem der Ansprüche 1 bis 9, dadurch gekennzeichnet, dass er anstelle des Gens, das für ein Protein/Produkt von Interesse kodiert, eine multiple Klonierungsstelle zum Einbau eines Gens, das für ein Protein von Interesse kodiert, umfasst.
11. Eukaryontische Wirtszelle, die mit einem Expressionsvektor nach einem der Ansprüche 1 bis 10 transfiziert worden ist.
- 15 12. Wirtszelle nach Anspruch 11, dadurch gekennzeichnet, dass es sich um eine Säugerzelle, vorzugsweise eine CHO-Zelle, handelt.
13. Wirtszelle nach Anspruch 11 oder 12, dadurch gekennzeichnet, dass sie zusätzlich mit einem oder mehreren Vektoren mit Genen, die für ein oder mehrere andere Proteine/Produkte von Interesse sowie zumindest einen weiteren Selektionsmarker kodieren, transfiziert worden ist.
- 20 14. Verfahren zur Herstellung eines heterologen Genprodukts, bei dem eine Wirtszelle nach einem der Ansprüche 11 bis 13 unter Bedingungen, die eine Expression des Genprodukts ermöglichen, kultiviert und das Genprodukt aus der Kultur oder dem Kulturmedium isoliert wird.
- 25 15. Verfahren zur Herstellung eines heteromeren Proteins/Produkts, bei dem eine Wirtszelle nach Anspruch 13, die mit Expressionsvektoren, welche für unterschiedliche Untereinheiten des heteromeren Proteins/Produkts kodieren, co-transfiziert worden ist, unter Bedingungen, die eine Expression des heteromeren Proteins/Produkts ermöglichen, kultiviert und das heteromere Protein/Produkt aus der Kultur oder dem Kulturmedium isoliert wird.

16. Verfahren nach Anspruch 15, dadurch gekennzeichnet, dass das heteromere Protein ein Antikörper ist.
- 5 17. Verfahren nach einem der Ansprüche 14 bis 16, dadurch gekennzeichnet, dass die Wirtszelle einem oder mehreren Genamplifikationsschritten unterzogen wird.
- 10 18. Verfahren nach Anspruch 17, dadurch gekennzeichnet, dass der amplifizierbare Selektionsmarker DHFR und das Amplifikationsmittel Methotrexat ist.
- 15 19. Verfahren nach Anspruch 18, dadurch gekennzeichnet, dass die Wirtszelle lediglich einem Genamplifikationsschritt mit Methotrexat unterzogen wird.
- 20 20. Verfahren nach einem der Ansprüche 14 bis 19, dadurch gekennzeichnet, dass es in einem serumfreien Kulturmedium durchgeführt wird.
21. Verfahren nach einem der Ansprüche 14 bis 20, dadurch gekennzeichnet, dass die Wirtszelle in Suspensionskultur kultiviert wird.
- 25 22. Verfahren zur Selektion einer Wirtszelle, die ein Protein/Produkt von Interesse exprimiert, bei dem eine Population von Wirtszellen nach einem der Ansprüche 11 bis 13 unter Bedingungen kultiviert wird, die eine Expression des Proteins/Produkts von Interesse und des fluoreszierenden Proteins ermöglichen, und die Zelle(n) identifiziert und/oder selektiert werden, welche den höchsten Expressionslevel an fluoreszierendem Protein zeigen.
- 30 23. Verfahren nach Anspruch 22, dadurch gekennzeichnet, dass die Selektion mit Hilfe eines Fluoreszenz-aktivierten Zellsortiergeräts (FACS) erfolgt.

24. Verfahren nach Anspruch 22 oder 23, dadurch gekennzeichnet, dass die entsprechend selektierten Wirtszellen einem oder mehreren zusätzlichen Genamplifikationsschritten unterzogen werden.
- 5 25. Verfahren nach Anspruch 24, dadurch gekennzeichnet, dass der amplifizierbare Selektionsmarker DHFR und das Amplifikationsmittel Methotrexat ist.

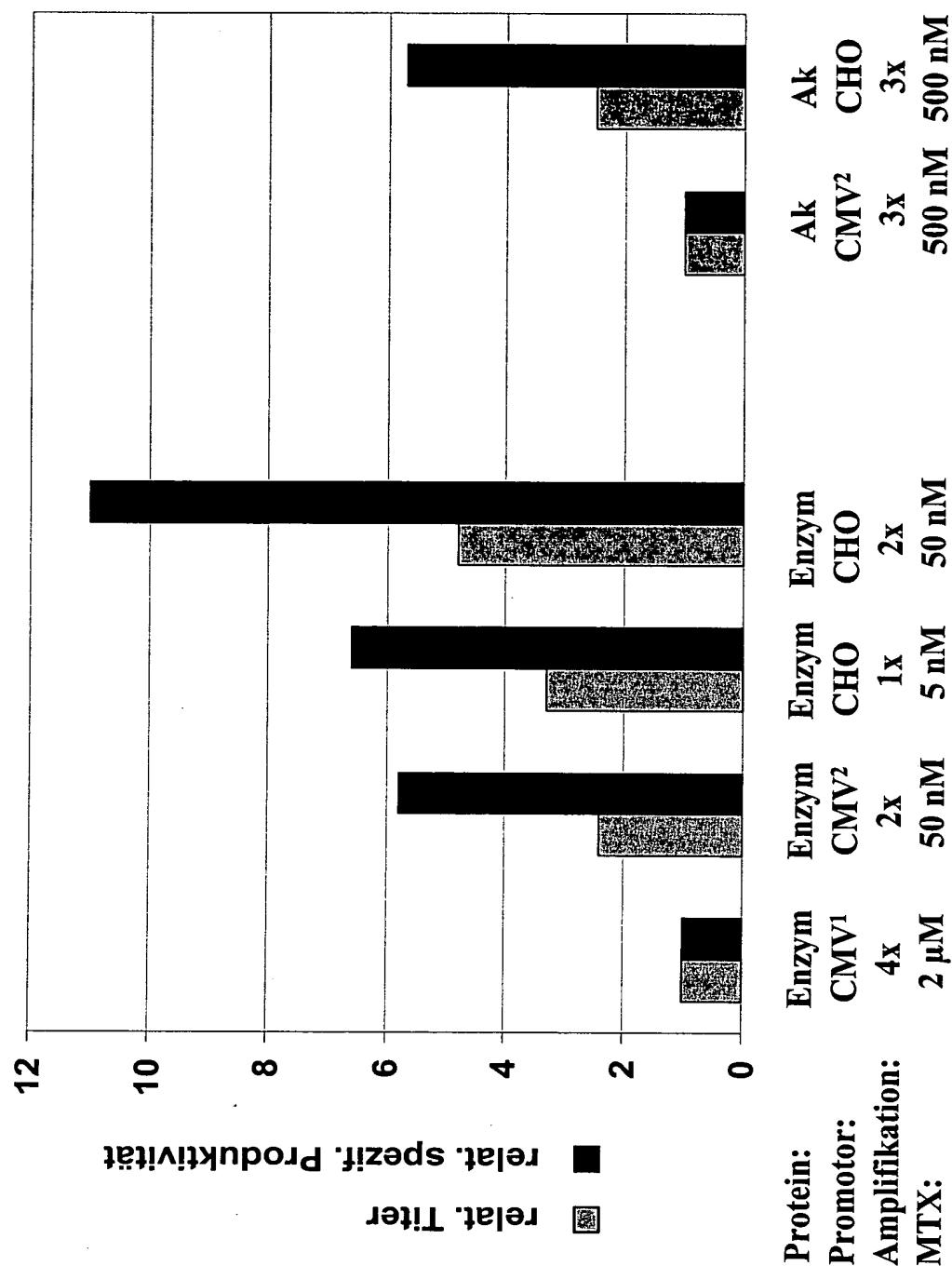
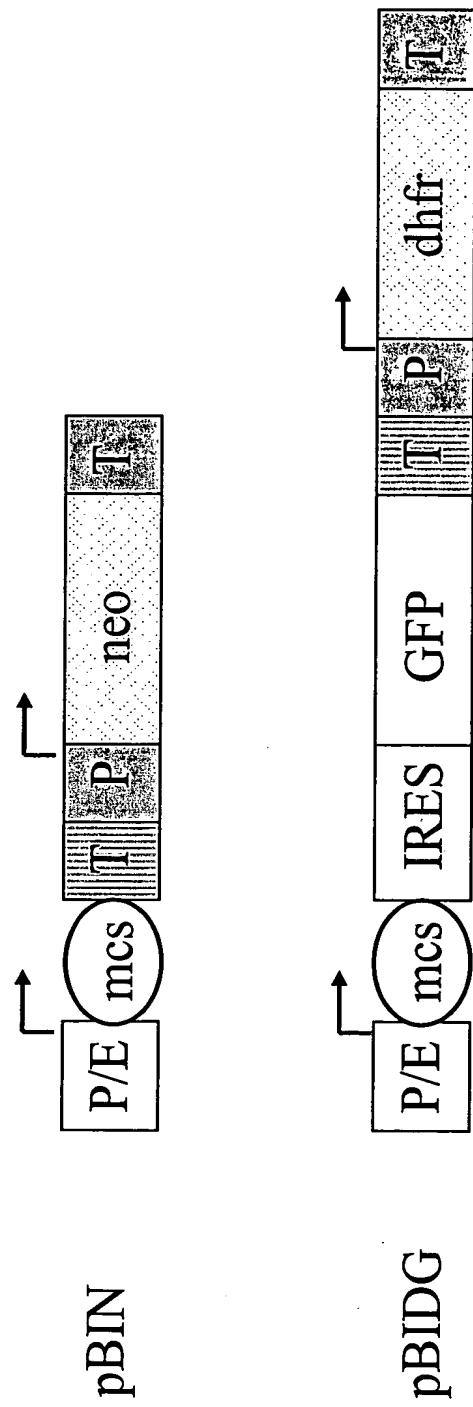
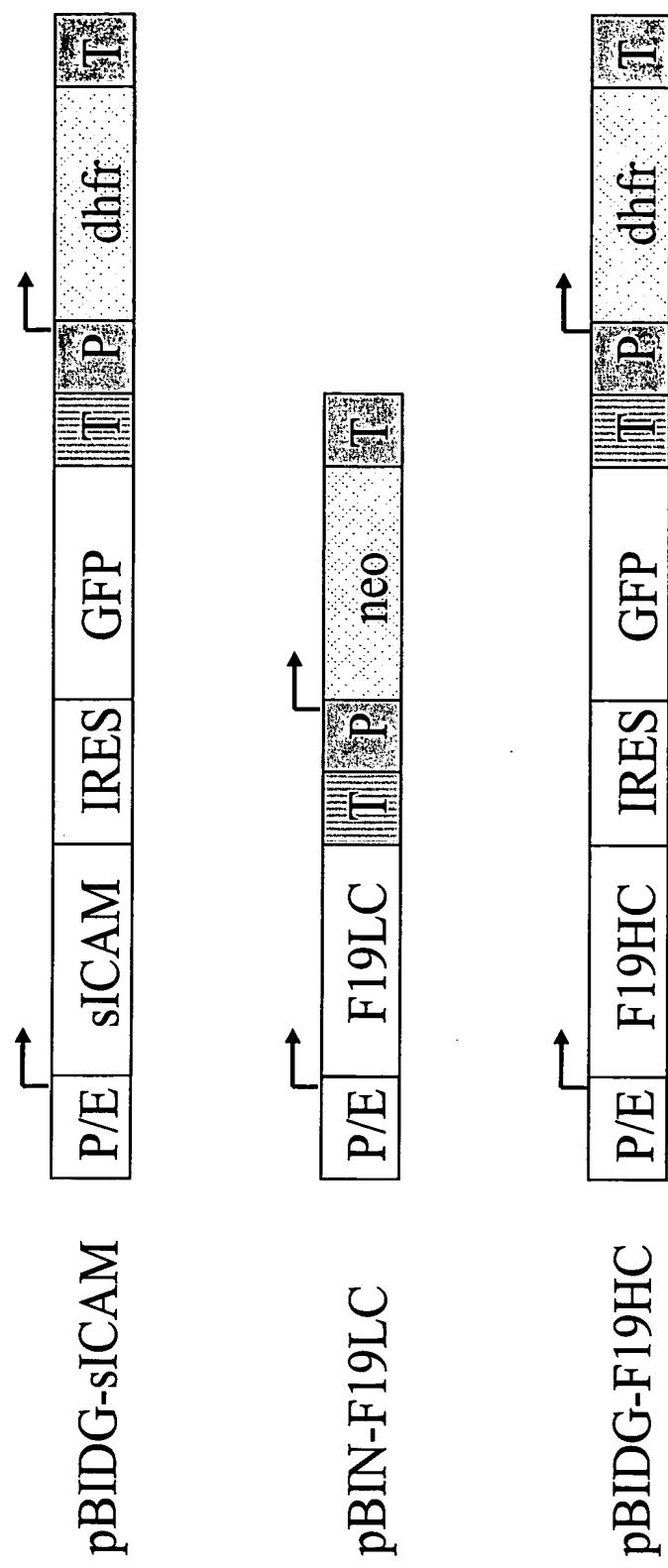


Figure 1

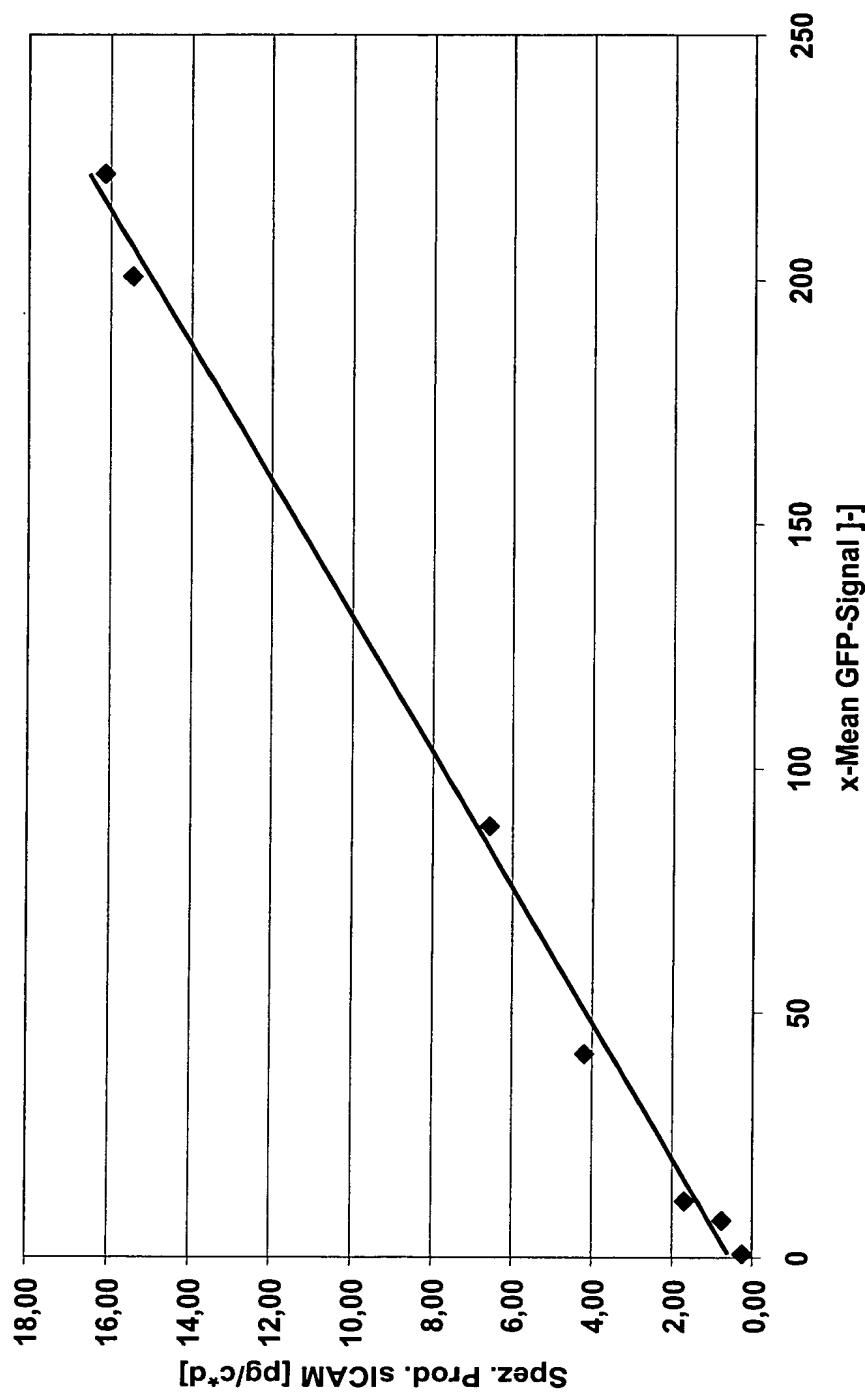


Figur 2



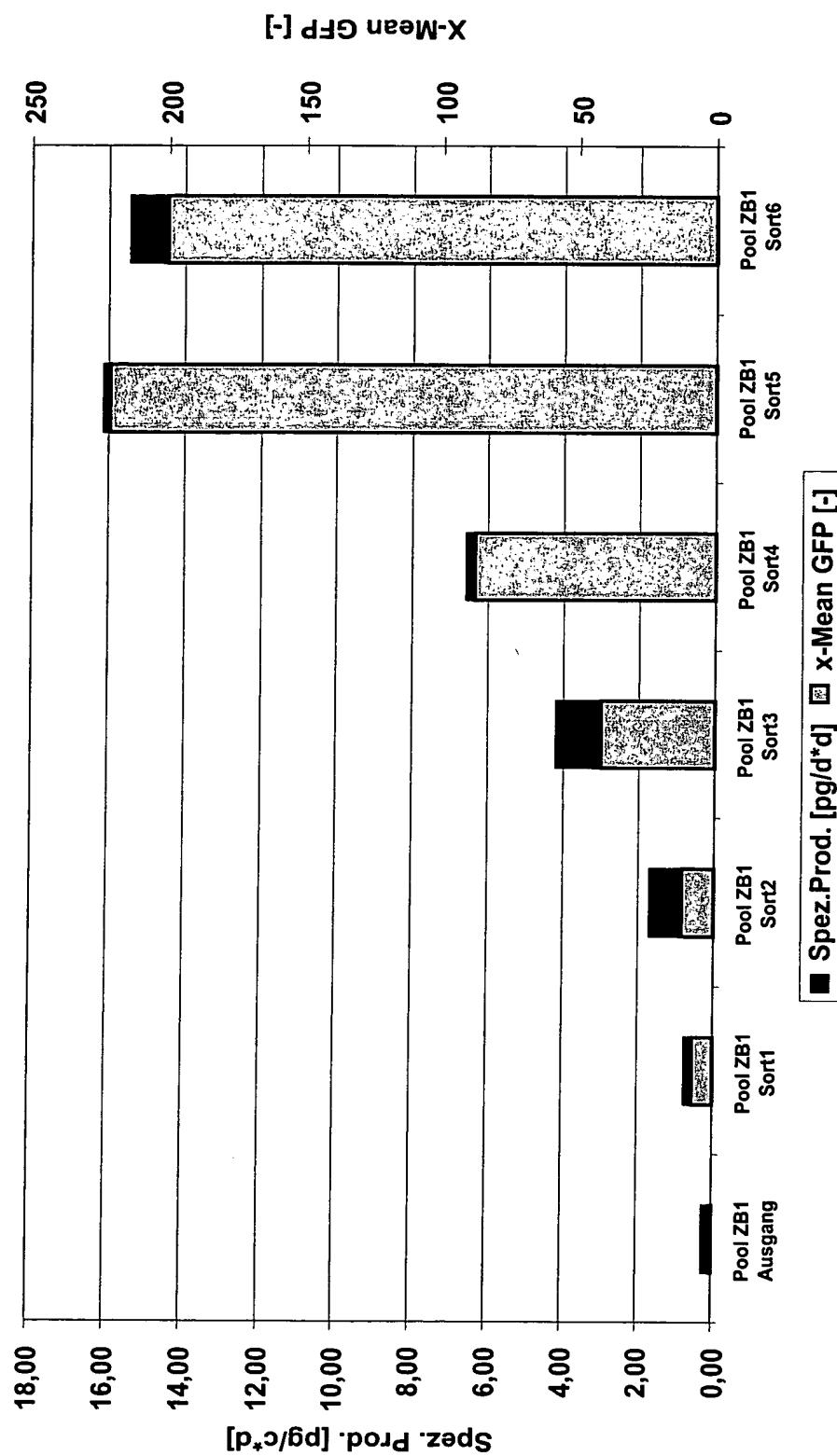
Figur 3

4/9



Figur 4

5/9



Figur 5

6/9

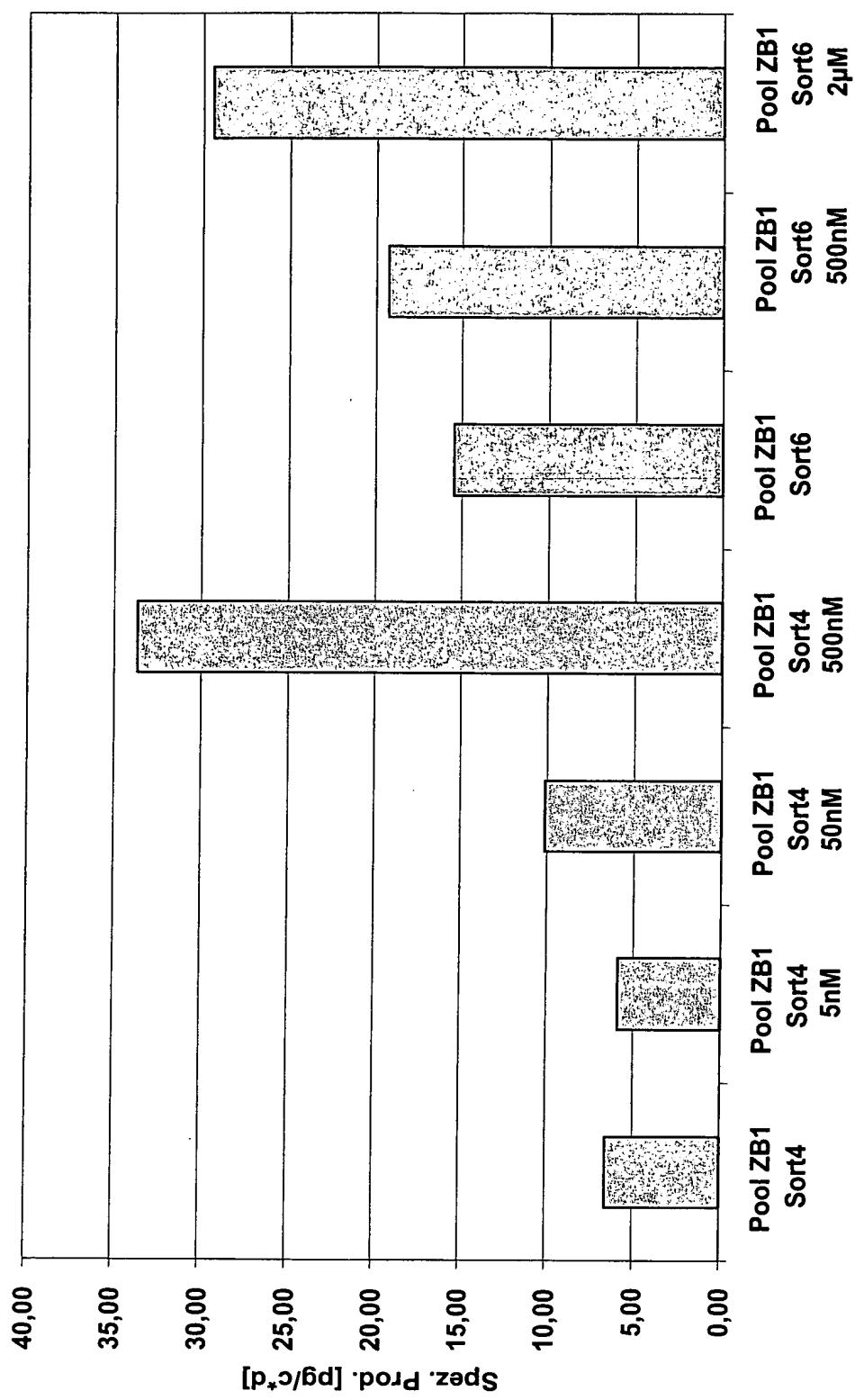


Figure 6

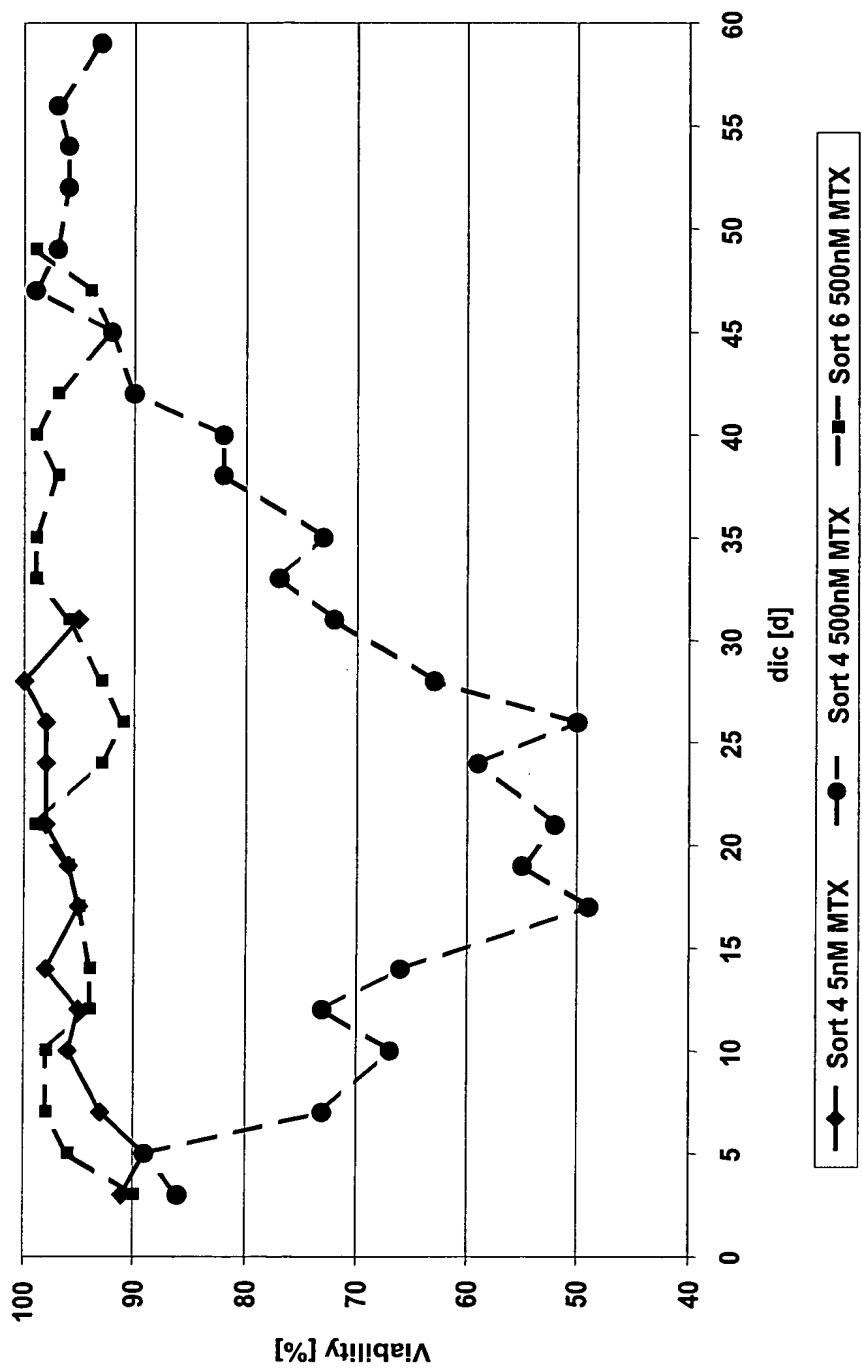


Figure 7

8/9

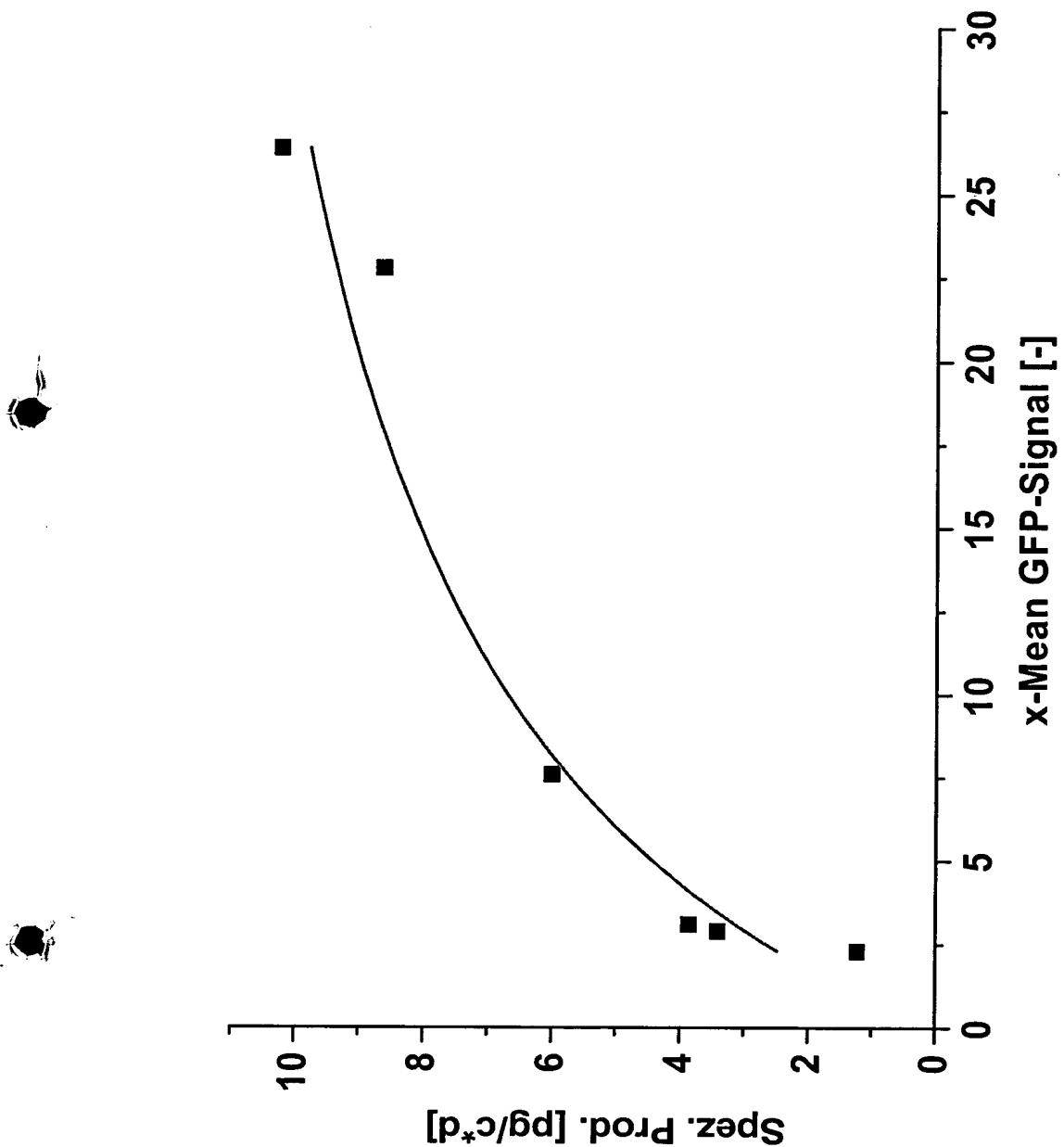


Figure 8

9/9

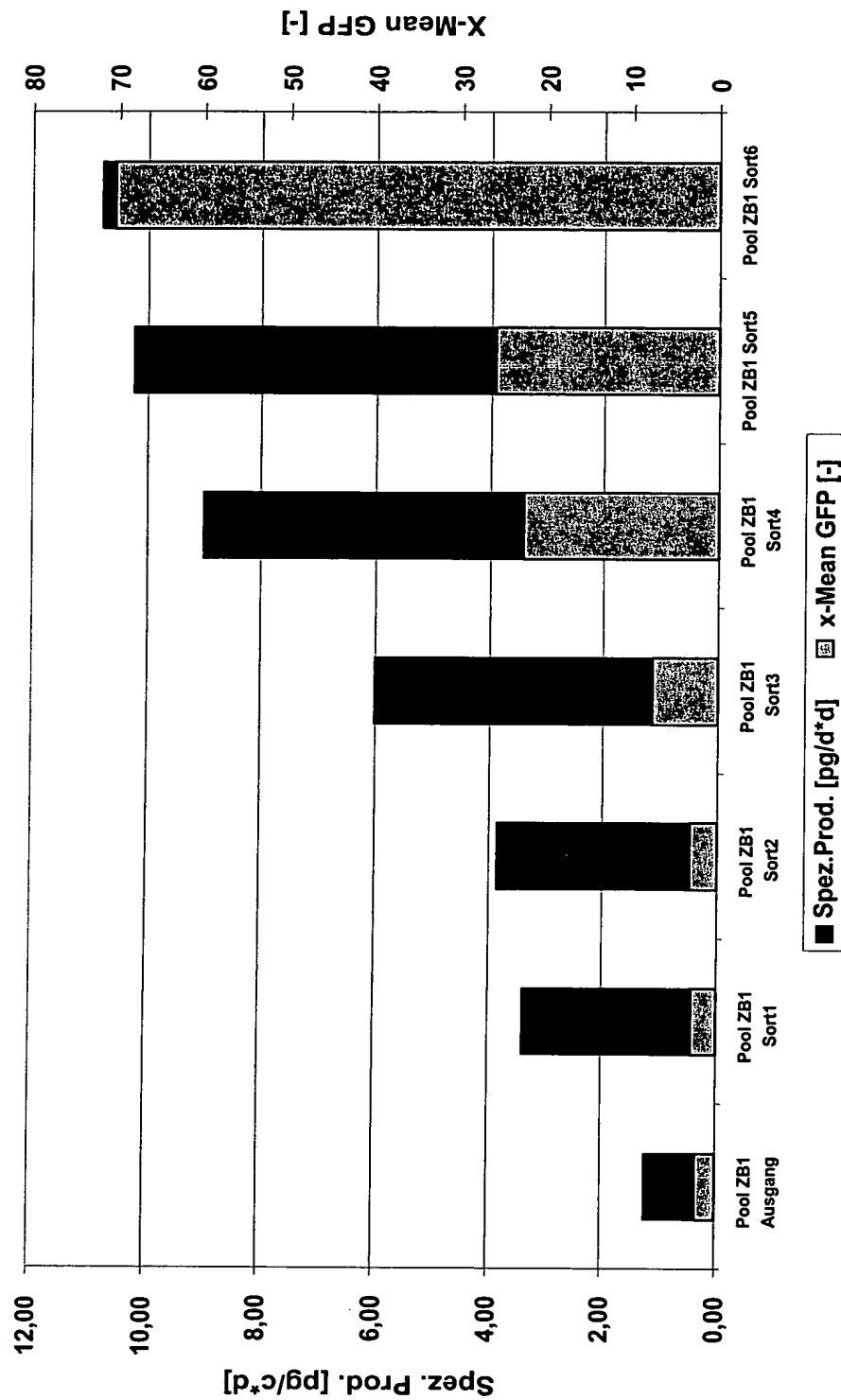


Figure 9

ZUSAMMENFASSUNG

5 EXPRESSIONSVEKTOR, VERFAHREN ZUR HERSTELLUNG VON
HETEROLOGEN GENPRODUKTEN UND SELEKTIONSVERFAHREN FÜR
HOCHPRODUZIERENDE REKOMBINANTE ZELLEN

Ein Expressionsvektor für eukaryontische Zellen umfassend ein Gen, das für ein Protein von Interesse kodiert, in funktioneller Verknüpfung mit einem Hamster-
10 Ubiquitin/S27a-Promotor und einem Gen, das für ein fluoreszierendes Protein kodiert. Vorzugsweise enthält der Expressionsvektor noch ein amplifizierbares
1 Selektionsmarkergen.

15 Beschrieben werden ferner mit dem Expressionsvektor transfizierte Wirtszellen, vorzugsweise Säugerzellen, Verfahren zur Herstellung heterologer Genprodukte und ein Verfahren zur Selektion hochproduzierender Zellen.